सांख्यिकी का परिचय

कक्षा 11 के लिए सांख्यिकी की पाठ्यपुस्तक

लेखक

ए.एल. नागर

संपादक

जी.एस. कुशवाहा

हिंदी रूपांतरण

नीरजा रश्मि

राष्ट्रीय अनुसंधान और प्रशिक्षण परिषद्

NATIONAL COUNCIL OF EDUCATIONAL RESEARCH AND TRAINING

प्रथम संस्करण
सितम्बर 2002
भाद्रपद 1924

ISBN 81-7450-084-7

PD 30T DRH

एन.सी.ई.आर.टी. के प्रकाशन विभाग के कार्यालय

एन.सी.ई.आर.टी. कैम्पस	108, 100 फीट रोड, होस्डेकेरे	नवजीवन ट्रस्ट भवन	सी.डब्लू.सी. कैम्पस
श्री अरविंद मार्ग	हेली एक्सटेंशन बनाशंकरी III इस्टेज	डाकघर नवजीवन	32, बी.टी. रोड, सुखचर
नई दिल्ली 110 016	**बैंगलूर 560 085**	**अहमदाबाद 380 014**	**24 परगना 743 179**

प्रकाशन सहयोग

संपादन : दयाराम हरितश
उत्पादन : प्रमोद रावत
राजेन्द्र चौहान

आवरण
डी.के. शिंदे

रु. 24.00

एन.सी.ई.आर.टी. वाटर मार्क 70 जी.एस.एम. पेपर पर मुद्रित।

प्रकाशन विभाग में सचिव, राष्ट्रीय शैक्षिक अनुसंधान और प्रशिक्षण परिषद्, श्री अरविंद मार्ग, नई दिल्ली 110 016 द्वारा प्रकाशित तथा एस.पी.ए. प्रिन्टर्स (प्रा.) लि., बी-17/3, ओखला इंडस्ट्रियल एरिया, फेज़-II, नई दिल्ली 110 020 द्वारा मुद्रित।

प्राक्कथन

प्रस्तुत पाठ्यपुस्तक **सांख्यिकी का परिचय** (कक्षा XI, प्रथम सेमेस्टर) उच्चतर माध्यमिक स्तर के लिए तैयार की गई अर्थशास्त्र की चार पाठ्यपुस्तकों की शृंखला में पहली है। यह 'विद्यालयी शिक्षा के लिए राष्ट्रीय पाठ्यचर्या की रूपरेखा, 2000' के अनुरूप तैयार किए गए अर्थशास्त्र के नए पाठ्यक्रम पर आधारित है।

माध्यमिक स्तर तक अर्थशास्त्र सामाजिक विज्ञान की समेकित पाठ्यचर्या के एक अंग के रूप में पढ़ाया जाता है। उच्चतर माध्यमिक स्तर पर पहली बार एक स्वतंत्र विषय के रूप में अर्थशास्त्र से विद्यार्थियों का परिचय होता है। उच्चतर माध्यमिक स्तर के दो वर्षों की अवधि छः माह के चार सेमेस्टरों में विभाजित होती है। इन चारों सेमेस्टरों के लिए अर्थशास्त्र में चार पाठ्यपुस्तकें हैं, जिनमें से दो ग्यारहवीं कक्षा के लिए हैं तथा दो बारहवीं कक्षा के लिए।

ग्यारहवीं कक्षा के प्रथम सेमेस्टर के लिए तैयार की गई इस पाठ्यपुस्तक में स्पष्ट उदाहरणों तथा आरेखी रूपायनों के द्वारा आधारभूत सांख्यिकी का परिचय देने का प्रयास किया गया है। विद्यार्थियों की आवश्यकताओं तथा योग्यता को ध्यान में रखते हुए, विवरणों को सरल तथा स्पष्ट रखा गया है। पुस्तक में प्रयुक्त अधिकांश आँकड़े और उदाहरण भारतीय परिप्रेक्ष्य में वास्तविक जगत की समस्याओं पर आधारित हैं। पुस्तक में परियोजना कार्य भी दिए गए हैं, जो विद्यार्थियों को विषय के सैद्धांतिक ज्ञान को व्यवहार में प्रयोग करने की कला सिखाने में उपयोगी होंगे।

इस पाठ्यपुस्तक को तैयार करने में परिषद् को कई स्थानों से सहायता मिली हैं। मैं लेखक महोदय के प्रति आभार व्यक्त करता हूँ जिन्होंने पुस्तक को बोधगम्य रूप में लिखने का कठिन कार्य कुशलता से संपन्न किया। विद्यालयों, महाविद्यालयों तथा विश्वविद्यालयों के उन शिक्षकों के प्रति भी आभार व्यक्त करता हूँ जिन्होंने विषय-विशेषज्ञों के रूप में पुस्तक की पांडुलिपि के पुनरीक्षण हेतु आयोजित कार्यशालाओं में भाग लेकर पुस्तक को वर्तमान स्वरूप प्रदान करने में सक्रिय योगदान दिया।

आशा है कि यह पाठ्यपुस्तक उन विद्यार्थियों के लिए उपयोगी होगी जिन्होंने एक स्वतंत्र विषय के रूप में अर्थशास्त्र का अध्ययन आरंभ किया है। पाठ्यचर्या एवं पाठ्यपुस्तकों का निर्माण एक अनवरत प्रक्रिया है, अतः भविष्य में पुस्तक में सुधार हेतु सुझावों का स्वागत किया जाएगा।

जगमोहन सिंह राजपूत
निदेशक
राष्ट्रीय शैक्षिक अनुसंधान और प्रशिक्षण परिषद्

जून 2002
नई दिल्ली

पाठ्यपुस्तक पुनरीक्षण समिति

1. ए.एल. नागर
   आचार्य, अर्थशास्त्र (अवकाश प्राप्त)
   दिल्ली स्कूल ऑफ इकॉनोमिक्स
   दिल्ली विश्वविद्यालय, दिल्ली

2. जी.एस. कुशवाहा
   आचार्य (अवकाश प्राप्त)
   अर्थशास्त्र विभाग
   बनारस हिन्दू विश्वविद्यालय, वाराणसी

3. ए.एस. गर्ग
   स्नातकोत्तर अध्यापक (अर्थशास्त्र)
   राजकीय प्रतिभा विकास विद्यालय
   सूरज मल विहार, दिल्ली

4. वीणा गुप्ता
   प्रवक्ता, अर्थशास्त्र
   राजकीय कन्या उ. मा. विद्यालय न.-1
   सरोजिनी नगर, नई दिल्ली

5. सी.के. बन्सल
   स्नातकोत्तर अध्यापक (अर्थशास्त्र)
   केंद्रीय विद्यालय
   विकास पुरी, नई दिल्ली

6. कान्ता जोशी
   स्नातकोत्तर अध्यापक (अर्थशास्त्र)
   राजकीय कन्या उ.मा.विद्यालय न.-2
   किदवई नगर, नई दिल्ली

7. के.सी. शर्मा
   उप-प्रधानाचार्य
   राजकीय को.एड. सर्वोदय विद्यालय
   शिकारपुर,नई दिल्ली

8. नीरजा रश्मि **(संयोजक)**
   प्रवाचक, अर्थशास्त्र
   सामाजिक विज्ञान एवं मानविकी शिक्षा विभाग
   एन.सी.ई.आर.टी., नई दिल्ली

विषय-सूची

गांधी जी का जन्तर

तुम्हें एक जन्तर देता हूं। जब भी तुम्हें सन्देह हो या तुम्हारा अहम् तुम पर हावी होने लगे, तो यह कसौटी आजमाओ :

जो सबसे गरीब और कमजोर आदमी तुमने देखा हो, उसकी शकल याद करो और अपने दिल से पूछो कि जो कदम उठाने का तुम विचार कर रहे हो, वह उस आदमी के लिए कितना उपयोगी होगा। क्या उससे उसे कुछ लाभ पहुंचेगा ? क्या उससे वह अपने ही जीवन और भाग्य पर कुछ काबू रख सकेगा ? यानि क्या उससे उन करोड़ों लोगों को स्वराज्य मिल सकेगा जिनके पेट भूखे हैं और आत्मा अतृप्त है ?

तब तुम देखोगे कि तुम्हारा सन्देह मिट रहा है और अहम् समाप्त होता जा रहा है।

अध्याय - 1

परिचय

1. सांख्यिकी का अर्थ

ऑक्सफोर्ड शब्दकोष के अनुसार सांख्यिकी शब्द का प्रयोग यदि बहुवचन में हो तो इसका अर्थ 'क्रमबद्ध एकत्रित संख्यात्मक तथ्य' होता है। एकवचन में इसका अर्थ 'आँकड़ों को एकत्रित करने या उनका प्रयोग करने का विज्ञान' या 'सांख्यिकीय तथ्य तथा सांख्यिकीय मदें' होता है।

अंग्रेजी भाषा की रैन्डम हाउस डिक्शनरी 'सांख्यिकी' का निम्नलिखित अर्थ बताती है: *एकवचन में 'सांख्यिकी एक ऐसा विज्ञान है जो संख्यात्मक तथ्यों व आँकड़ों के एकत्रीकरण, वर्गीकरण, विश्लेषण तथा प्रतिपादन से संबंध रखता है तथा प्रायिकता (probability) के गणितीय सिद्धांतों का प्रयोग कर, लगभग पृथक तत्वों वाले समुच्चयों को एक व्यवस्था तथा नियमितता प्रदान करता है।'* बहुवचन में *'सांख्यिकी का अभिप्राय संख्यात्मक तथ्यों तथा आँकड़ों'* से है।

सांख्यिकीय आँकड़ों के विश्लेषण की विधियाँ विवरणात्मक या प्रायिकतात्मक हो सकती हैं। विवरणात्मक विधि में निम्नलिखित शामिल हैं :

क. आँकड़ों का एकत्रीकरण व वर्गीकरण,

ख. आँकड़ों का सारणीयन एवं आरेखी प्रस्तुतीकरण,

ग. आँकड़ों की विशेषताएँ सूचित करने वाले सूचकों का आकलन।

आगे के अध्यायों में हम इनकी चर्चा विस्तार से करेंगे। प्रायिकतात्मक विधियों की चर्चा इस पुस्तक की विषयवस्तु में नहीं है। सांख्यिकीय आँकड़े मात्रात्मक या गुणात्मक हो सकते हैं।

मात्रात्मक आँकड़े (Quantitative Data)

यदि चरों को संख्याओं द्वारा मापा जा सकता हो तो इन्हें मात्रात्मक आँकड़े कहते हैं। उदाहरण के लिए, दैनिक तापमान, व्यक्तियों के कद व वजन, कीमतें व आय इत्यादि मात्रात्मक चर हैं। इनके मूल्य संख्याओं के द्वारा दिखाए जा सकते हैं।

गुणात्मक आँकड़े (Qualitative Data)

कभी-कभी चरों को उस प्रकार संख्याओं द्वारा मापना संभव नहीं होता जैसे हम व्यक्तियों की लंबाई या वजन को या वस्तुओं की कीमतों को या व्यक्तियों की आय को मापते हैं। उदाहरण के लिए, राजनैतिक व्यवस्था के बारे में लोगों के विचार, उनकी बुद्धिमत्ता, संगीत व कला के प्रति उनकी अभिरुचि, व्यक्तियों या वस्तुओं (जैसे फूलों, बाग-बगीचों आदि) की सुंदरता आदि को संख्यात्मक रूप में मापना संभव नहीं। उन्हें, उनके गुणों के अनुसार, कोटियों में बाँटा जा सकता है। हम वस्तुओं की तुलना कर यह कह सकते हैं कि क, ख की तुलना में अधिक सुंदर है ; ख, ग की तुलना में अधिक सुंदर है आदि। सबसे सुंदर व्यक्ति को पहली कोटि में, दूसरे को दूसरी कोटि में रखा जा सकता है। सबसे कम सुंदर को निम्नतम कोटि में रखा जा

सकता है। इसी प्रकार, कलाकारों (संगीतकारों, पेंटरों आदि) के प्रदर्शनों को संख्यात्मक रूप से नहीं मापा जा सकता है, परंतु निर्णायकगण इन्हें भिन्न-भिन्न कोटियों में रख सकते हैं। सांख्यिकीय विश्लेषण में कोटियों का प्रयोग संख्यात्मक मापन के स्थान पर किया जा सकता है।

2. सांख्यिकी की विषयवस्तु

सांख्यिकीय आँकड़ों का प्रयोग हम दैनिक जीवन में करते हैं। उदाहरण के लिए, एक महीने का पारिवारिक बजट बनाने के लिए विभिन्न वस्तुओं तथा सेवाओं की कीमतों के आँकड़ों की आवश्यकता पड़ती है, ताकि यह निश्चित किया जा सके कि उपभोग के विभिन्न मदों (खाद्य-सामग्रियों, वस्त्र, यात्रा, बच्चों की पढ़ाई इत्यादि) पर आय का कितना अनुपात आबंटित किया जाए? परीक्षा में विद्यार्थियों का निष्पादन उनके द्वारा प्राप्त किए गए अंकों पर निर्भर करता है। विभिन्न स्कूलों के बीच तुलना दसवें वर्ग में उत्तीर्ण विद्यार्थियों की संख्या के द्वारा की जा सकती है अर्थात् यह जानकर कि कितने विद्यार्थी प्रथम, द्वितीय तथा तृतीय श्रेणी में उत्तीर्ण हुए हैं और कितने अनुत्तीर्ण।

अर्थशास्त्री सांख्यिकीय आँकड़ों का प्रयोग विभिन्न वस्तुओं और सेवाओं की कीमतों की प्रवृत्ति का विश्लेषण करने के लिए तथा अर्थव्यवस्था में उपभोग व उत्पादन के स्वरूपों का विश्लेषण करने के लिए करते हैं।

सरकार आँकड़े एकत्र करने वाली सबसे बड़ी संस्था है, जो जनांकिकीय विशेषताओं (जैसे जन्म व मृत्यु-दर, जनसंख्या का आकार व संघटन आदि), राष्ट्रीय आय, उपभोग, औद्योगिक तथा कृषि उत्पादन आदि के संबंध में आँकड़े एकत्र करती है।

सरकार तथा नीति निर्धारक, सांख्यिकीय आँकड़ों का प्रयोग आर्थिक विकास के लिए उपयुक्त नीतियों आदि के निर्माण में करते हैं।

व्यवसाय में, सांख्यिकीय आँकड़ों का प्रयोग ब्रिक्री तथा कीमतों के संबंध का अध्ययन करने के लिए तथा बाजार के उतार-चढ़ाव को निश्चित करने के लिए किया जाता है। सांख्यिकीय आँकड़ों का प्रयोग किसी दिए गए समय-बिंदु पर स्टॉक की मात्राओं के निर्धारण के लिए भी किया जा सकता है।

उद्योग में, आगतों तथा निर्गतों के संबंध के विश्लेषण में भी इसकी उपयोगिता है।

सांख्यिकी की प्रयोगात्मक तथा अप्रयोगात्मक विज्ञानों में भी महत्त्वपूर्ण भूमिका है। दोनों ही स्थितियों में चरों के बीच सांख्यिकीय संबंध प्रेक्षित आँकड़ों (observed data) के आधार पर प्राप्त किए जाते हैं।

भौतिकी, रसायन-शास्त्र जैसे प्रयोगात्मक विज्ञानों में प्रयोगशालाओं में सूक्ष्म उपकरणों के प्रयोग द्वारा वैज्ञानिक स्वयं आँकड़े तैयार करता है। अतः तुलनात्मक दृष्टि से प्रायोगिक अनुसंधान करने वाले वैज्ञानिकों के आँकड़े अप्रायोगिक अनुसंधान करने वाले वैज्ञानिकों (जैसे अर्थशास्त्रियों) के आँकड़ों से अधिक सही होते हैं। अर्थशास्त्रियों के आँकड़े दूसरों के द्वारा एकत्रित व संसाधित (processed), वर्गीकृत तथा सारणीबद्ध किए जाते हैं। कभी-कभी अर्थशास्त्री अपने लिए आँकड़े स्वयं एकत्र करते हैं जो प्राथमिक आँकड़े (primary data) कहे जाते हैं।

चरों के बीच सांख्यिकीय संबंध सही नहीं भी होते हैं तथा उनमें त्रुटियाँ (errors) भी हो सकती हैं। आँकड़ों के एकत्रीकरण, संगठन, वर्गीकरण, सारणीयन या प्रस्तुतीकरण की प्रक्रिया में भी कई प्रकार की त्रुटियाँ हो सकती हैं। इन त्रुटियों की चर्चा हम अगले अध्यायों में करेंगे। सांख्यिकीय संबंधों में छोड़े गए चरों के कारण त्रुटियाँ होती हैं। उदाहरण के लिए, किसी वस्तु की माँग की गई मात्रा तथा इसकी कीमत के संबंध को लें। हम जानते हैं कि कीमतों के बढ़ने पर माँग की गई

मात्रा में कमी आती है। परंतु, आय में परिवर्तन के कारण भी माँग की गई मात्रा में परिवर्तन होता है। यदि हम माँग की गई मात्रा तथा कीमत के संबंध को ही ध्यान में रखें (अन्य चरों के प्रभाव पर ध्यान न देते हुए), तो इसमें त्रुटियाँ अनिवार्यतः होंगी। इस संबंध में त्रुटि, छूटे हुए चरों के कारण होती है।

सांख्यिकीय विधियाँ, त्रुटियों की उपस्थिति में (जो आँकड़ों के मापन में त्रुटियों तथा छोड़े गए चरों के कारण होती हैं), सांख्यिकीय संबंधों के मापन में सहायक होती हैं। वस्तुतः सांख्यिकीय संबंधों में त्रुटियों की उपस्थिति ही (प्रयोगात्मक तथा अप्रयोगात्मक विज्ञान दोनों में) सांख्यिकीय विधियों को अपरिहार्य बना देती है। यदि त्रुटियाँ न होतीं तो सांख्यिकीय विधियों के प्रयोग की आवश्यकता पड़ती ही नहीं तथा गणितीय विधियाँ ही पर्याप्त होतीं।

3. अर्थशास्त्र में सांख्यिकी का महत्त्व

अर्थशास्त्र में वस्तुतः हमें कुछ चरों में परिवर्तन को मापने की आवश्यकता होती है, यदि अन्य चरों में परिवर्तन हो तो। उदाहरण के लिए,किसी वस्तु की कीमत में एक प्रतिशत की वृद्धि हो तो इसकी माँग में कितनी कमी होगी? दूसरा उदाहरण लें यदि भारतीय रिजर्व बैंक ने ब्याज की दर में आधे प्रतिशत की कमी की, तो अर्थव्यवस्था में निवल निजी निवेश में कितनी वृद्धि होगी? किसी फर्म के उत्पादन में कितना परिवर्तन होगा यदि प्रबंधन ने श्रमिकों की संख्या में वृद्धि करने का या प्रौद्योगिकी में अधिक धन निवेश करने का निर्णय किया?

सामान्यतः यह कहा जा सकता है कि किसी वस्तु की कीमत बढ़ने पर उसकी माँग की गई मात्रा घटती है। परंतु माँग की मात्रा में आनुपातिक परिवर्तन इस बात पर निर्भर करता है कि क्या यह वस्तु एक अनिवार्य वस्तु है (जैसे नमक या चावल आदि) या एक विलासित वस्तु (जैसे टी.वी. या रेफ्रिजरेटर)।

यदि दिए गए आँकड़ों के एक समुच्चय के लिए माँग की गई मात्रा तथा सांख्यिकीय संबंधों को मापा जा सके तो इसका प्रयोग माँग की कीमत लोच के संख्यात्मक माप को प्राप्त करने के लिए हो सकता है। यदि संख्यात्मक लोच (जिसे कीमत में आनुपातिक परिवर्तन के कारण माँग-मात्रा में सापेक्ष परिवर्तन के रूप में परिभाषित किया जा सकता है) शून्य के निकट है, तो कहा जा सकता है कि कीमत में थोड़े परिवर्तन के कारण माँग-मात्रा में भी थोड़ा ही परिवर्तन होगा। अन्यथा, यदि लोच अधिक है (1 से अधिक) तो कीमत में कम परिवर्तन का माँग-मात्रा पर बहुत अधित असर होगा। हम अनुमानित माँग-कीमत संबंध का प्रयोग कर, दी गई कीमतों के अनुरूप माँग की मात्रा को जान सकते हैं।

अर्थशास्त्र में **एकल समीकरण निदर्श** (Single Equation Models) भी हैं जो क्रेता, विक्रेता, निवेशकर्ता आदि के व्यवहारों का वर्णन करते हैं। ये सामान्यतः यथातथ (exact) समीकरण नहीं होते हैं (इसका कारण छोड़े गए चर तथा चरों के मापन में त्रुटियाँ हैं जैसा पहले बताया जा चुका है)।

बहु-समीकरण निदर्श (Multi Equation Models) का प्रयोग बाजार की संरचना (जहाँ विक्रेताओं व क्रेताओं का संपर्क होता है), अर्थव्यवस्था की संरचना (जहाँ कई आर्थिक एजेंटों, जैसे, उपभोक्ता, उत्पादक, श्रमिक आदि का आपस में संपर्क होता है) आदि का वर्णन करने के लिए होता है। अर्थव्यवस्था के विभिन्न क्षेत्रकों का वर्णन करने के लिए निदर्श उपलब्ध हैं जैसे, कृषि, उद्योग तथा निवेश-व्यापार के निदर्श।

4. सांख्यिकीय विधियों का दुरुपयोग

सभी विज्ञानों की भाँति सांख्यिकीय विधियाँ भी सुपरिभाषित मान्यताओं पर आधारित हैं। सांख्यिकीय परिणामों के विश्लेषण तथा प्रतिपादन के लिए उन

मान्यताओं का पूरा ध्यान रखा जाना चाहिए, जिन पर वे आधारित हों। यदि उन्हें उपेक्षित किया गया तो परिणाम गलत होंगे।

सांख्यिकीय विश्लेषण में आँकड़ों की गुणवत्ता बहुत महत्त्वपूर्ण है। त्रुटियों की संभावना तभी होती है जब ऐसे आँकड़ों का प्रयोग किया जा रहा हो जो या तो अपर्याप्त हों, या, आँकड़ों का प्रतिनिधित्व न करते हों, या आपस में तुलना करने लायक न हों। उदाहरण के लिए, मान लें किसी दवा या खाद के प्रभाव की जाँच करनी है। इसके लिए इसे विभिन्न स्थितियों में जाँचा जाना चाहिए। यदि कुछ ही स्थितियों में इसके प्रभाव की जाँच कर कोई निर्णय ले लिया जाए, तो गलतियों की संभावना रहती है। इसी प्रकार, यदि किसी शहर में रहने वाले व्यक्तियों की औसत आय का अनुमान लगाना हो और अनुमान का आधार केवल धनी वर्ग के लोगों की आय को बनाया गया तो हमारा अनुमान वास्तविकता से अधिक होगा।

आँकड़ों के एकत्रीकरण में परिभाषाओं तथा अवधारणाओं में अस्पष्टता के कारण गलत परिणामों पर पहुँचने की संभावना रहती है। उदाहरण के लिए, यदि परिवारों की आय के आँकड़े एकत्र किए जा रहे हों तो सबसे पहले यह बताना जरूरी है कि हम आय का आकलन किस प्रकार करना चाहते हैं। जिन लोगों की मजदूरी या वेतन निश्चित है, वे तो कोई सही उत्तर दे सकते हैं। परंतु व्यापारियों की निश्चित आय नहीं होती है। इनकी आय हर महीने बदलती रहती है तथा ये अपनी आय का नियमित रूप से उचित लेखा नहीं रख पाते, अतः ये अपनी आय का सही आकलन कैसे कर सकते हैं? इसी प्रकार, मान लें कि हम रोजगार से संबद्ध आँकड़े एकत्र कर रहे हैं। कुछ लोग ऐसे हो सकते हैं जो दिन या माह के कुछ भाग में ही कार्य करते हैं। हम ऐसे लोगों के रोजगार का मापन किस प्रकार कर सकते हैं?

सांख्यिकीय विधियाँ सामान्य बुद्धि का स्थानापन्न नहीं हैं। अतः इनका प्रयोग बिना सोचे-समझे नहीं किया जाना चाहिए।

एक बार किसी राज्य के एक गाँव में महामारी फैल गई। सरकार ने महामारी को रोकने के उपाय शुरू कर दिए। दवाओं तथा डॉक्टरों की सहायता दी गई। एक नेता ने, जो सांख्यिकी का ज्ञान रखने का दावा भी करते थे, आँकड़े एकत्र किए तथा पाया कि उन गाँवों में मौतों की संख्या अधिक थी, जहाँ डॉक्टरों की संख्या अधिक थी। अतः उन्होंने यह निष्कर्ष निकाला कि मौतों के जिम्मेदार डॉक्टर थे, अतः उन्हें सजा होनी चाहिए।

दूसरी रोचक कहानी सामान्यतः सांख्यिकी का मजाक बनाने के लिए कही जाती है। कहा जाता है कि चार व्यक्तियों का एक परिवार (पति-पत्नी तथा दो बच्चे) किसी नदी को पार करने गए। पिता को नदी की औसत गहराई मालूम थी। अतः उसने अपने परिवार के सदस्यों की औसत लंबाई का हिसाब लगाया। चूँकि परिवार के सदस्यों की औसत लंबाई पानी की औसत गहराई से अधिक थी, उसने सोचा कि वे लोग नदी को सुरक्षित पार कर लेंगे। परिणामस्वरूप, नदी पार करते समय बच्चे पानी में बह गए। यहाँ क्या गलती हुई थी? गलती औसत निकालने के तरीके में थी या औसतों के दुरुपयोग में?

कई बार लोग 'सांख्यिकीय भ्रांतियों', 'सांख्यिकीय असत्यों' आदि की बात करते हैं। ऐसा आँकड़ों को गलत ढंग से संचालित करने या मूल मान्यताओं पर ध्यान न देने या कुछ सूचकों (जैसे औसतों) के दुरुपयोग के कारण होता है, जैसा ऊपर बताया गया है।

अभ्यास

1. 'सांख्यिकी को संख्यात्मक तथ्यों के समुच्चय के रूप में परिभाषित करते हैं' कुछ उदाहरण दें।
2. "सांख्यिकी ऐसा 'विज्ञान' है जो सांख्यिकीय आँकड़ों के विश्लेषण से संबंध रखता है।" उदाहरण दें।
3. उदाहरणों की सहायता से 'मात्रात्मक' तथा 'संख्यात्मक' आँकड़ों में अंतर बताएँ।
4. आँकड़ों को एकत्र करने की प्रक्रिया में किस प्रकार की त्रुटियों की संभावना रहती है?
5. सांख्यिकीय संबंध सामान्यतः यथातथ नहीं होते, क्यों?
6. प्रयोगात्मक तथा अप्रयोगात्मक विज्ञानों दोनों में सांख्यिकीय विधियाँ अपरिहार्य हैं - क्यों?
7. गणितीय तथा सांख्यिकीय संबंधों की तुलना करें तथा दोनों के उदाहरण दें।
8. 'सांख्यिकीय विधियाँ सामान्य बुद्धि का स्थानापन्न नहीं हैं।' व्याख्या करें।

अध्याय - 2

आँकड़ों का संकलन

यदि हम किसी आर्थिक या सामाजिक समस्या का अध्ययन करना चाहते हैं, तो हमें कुछ चरों से संबद्ध आँकड़ों की आवश्यकता होती है। उदाहरण के लिए यदि हम देखना चाहते हैं कि किसी वस्तु की माँग उसकी कीमत में परिवर्तन के साथ कैसे बदलती है, तो इसके लिए कई बाजारों में भिन्न-भिन्न समय-बिंदुओं पर उस वस्तु की माँग की गई मात्रा तथा उसकी कीमत के आँकड़ों की आवश्यकता होती है। यदि हम यह जानना चाहें कि कोई परिवार अपनी आय का कितना हिस्सा खाद्य-सामग्रियों, वस्त्र, बच्चों के स्कूल, मकान-किराया आदि पर खर्च करता है तो इसके लिए हमें परिवार के बजट से संबद्ध आँकड़ों की आवश्यकता पड़ती है।

हम बाजार-सर्वेक्षण द्वारा या पूछताछ द्वारा पारिवारिक बजट की जानकारी लेकर आँकड़ों का संग्रह कर सकते हैं या प्रकाशित आँकड़ों का प्रयोग कर सकते हैं अथवा ऑफिस रेकार्डों के द्वारा प्रदत्त आँकड़े भी लिए जा सकते हैं।

1. प्राथमिक तथा द्वितीयक आँकड़े (Primary and Secondary Data)

सांख्यिकीय अन्वेषण, या क्षेत्र-सर्वेक्षण द्वारा संग्रहीत आँकड़े **प्राथमिक आँकड़े** कहलाते हैं। इससे कोई फर्क नहीं पड़ता कि आँकड़े कौन इकट्ठे करता है। महत्त्वपूर्ण यह है कि वे आँखों देखी सूचना पर आधारित हैं। अन्वेषणकर्ता द्वारा स्वयं संग्रह किए गए आँकड़े प्राथमिक आँकड़े कहलाते हैं। अन्य एजेन्सियों (जैसे राष्ट्रीय प्रतिदर्श सर्वेक्षण संगठन, या भारतीय रिजर्व बैंक, आदि) द्वारा संग्रह किए गए आँकड़े भी यदि अन्वेषणकर्ता को मूल रूप में (प्रश्नावली या अनुसूची के रूप में) दिए गए, तो इसे भी प्राथमिक आँकड़े कहते हैं।

जो आँकड़े अन्वेषणकर्ता के द्वारा नहीं बल्कि किसी अन्य के द्वारा संग्रहीत तथा संसाधित (परीक्षण, सारणीकरण तथा आरेखी प्रस्तुतीकरण) किए गए हों, उन्हें द्वितीयक आँकड़े कहते हैं। अतः प्रकाशित आँकड़े अनिवार्यतः द्वितीयक आँकड़े होते हैं।

2. प्राथमिक आँकड़ों का संग्रह

जो व्यक्ति सांख्यिकीय अन्वेषण की योजना बनाता है तथा इसे संचालित करता है, उसे **अन्वेषणकर्ता** (investigator) कहते हैं। जो वास्तविक आँकड़ों का संग्रह करने क्षेत्र में जाता है उसे **गणनाकार** (enumerator) कहते हैं। **उत्तर-दाता** (respondents) प्रश्नावली के प्रश्नों का उत्तर देकर वास्तविक आँकड़े प्रदान करते हैं।

प्राथमिक आँकड़े निम्नलिखित विधियों से एकत्र किए जाते हैं:

i. जनगणना विधि (Census method) द्वारा, जहाँ दिए गए क्षेत्र की सभी व्यक्तिगत इकाइयाँ शामिल हों, या

ii. प्रतिदर्श विधि (sample method) द्वारा, जहाँ दिए गए क्षेत्र की जनसंख्या का केवल एक भाग शामिल हो।

क्षेत्रीय कार्य की योजना

आँकड़ों के संग्रह के लिए चाहे गणना विधि अपनाई जाए या प्रतिदर्श विधि, क्षेत्रीय कार्यों की योजना तथा संचालन सावधानी पूर्वक किया जाना चाहिए।

क. प्रश्नावली विधि

प्रश्नावली तैयार करने के लिए निम्नलिखित बातों का ध्यान रखना चाहिए :

i. **प्रश्नावली बहुत लंबी न हो :** लंबी प्रश्नावली से अन्वेषणकर्त्ता तथा उत्तरदाता दोनों ही थक जाते हैं।

ii. **प्रश्नावली के प्रश्न स्पष्ट तथा छोटे हों :** अस्पष्ट प्रश्न नहीं पूछे जाने चाहिएं क्योंकि वे गलत सूचना दे सकते हैं।

iii. **प्रश्नावली का निर्माण इस प्रकार किया जाना चाहिए कि प्रश्नावली के दूसरे प्रश्नों के साथ उनका अनुयोग (Cross checking) किया जा सके।**

iv. **प्रश्न ऐसे होने चाहिए कि गणनाकार या उत्तरदाता दोनों के लिए अधिक गणितीय हिसाब-किताब की आवश्यकता ना पड़े।**

ख. पूछ-ताछ की विधि

इस विधि के अंतर्गत आँकड़े निम्नलिखित दो प्रकार से एकत्र किए जा सकते हैं :

i. **साक्षात्कार विधि** द्वारा, अथवा

ii. **डाक प्रश्नावली विधि** द्वारा।

साक्षात्कार विधि के गुण

i. गणनाकार स्वयं ही उत्तरदाता को पूछताछ के उद्देश्य तथा अध्ययन का महत्त्व समझा सकता है।

ii. इससे उत्तरदाता से बेहतर सहयोग मिलने तथा प्रश्नावली के प्रश्नों के सही उत्तर मिलने में आसानी होती है।

iii. गणनाकार प्रश्नों को सही ढंग से समझने में तथा सही ढंग से उनके उत्तर देने में उत्तरदाताओं की मदद कर सकता है।

iv. इससे उत्तरदाता के समय की बचत होती है तथा वह प्रसन्न रहता है।

साक्षात्कार विधि के दोष

यह विधि खर्चीली है। इसके लिए हमें गणनाकारों की एक बड़ी टीम की आवश्यकता होती है। अतः उनके प्रशिक्षण तथा यात्रा पर खर्च होने के अतिरिक्त खाने-पीने, स्टेशनरी, रहने आदि पर भी खर्च होता है।

डाक प्रश्नावली विधि के गुण

डाक द्वारा उत्तरदाताओं को प्रश्नावली भेजा जाना अधिक सुविधाजनक तथा कम खर्चीला है।

डाक प्रश्नावली के दोष

i. उत्तरदाता प्रश्नों को सही ढंग से नहीं भी समझ सकते हैं या वे कुछ प्रश्नों का गलत अर्थ निकाल सकते हैं।

ii. उत्तरदाता सभी प्रश्नों का सही ढंग से उत्तर देने में असावधानी बरत सकते हैं।

iii. उत्तरदाता प्रश्नावली की उपेक्षा कर उन्हें कई बार न भी लौटा सकते हैं।

iv. कुछ प्रश्नावलियों के डाक में खो जाने का भी डर रहता है।

डाक प्रश्नावली विधि, उस स्थिति में उचित है जब उत्तरदाता के लिए कानूनी तौर पर सूचना देना अनिवार्य होता है। उदाहरण के लिए राजकीय संस्थाओं, बैंक तथा कम्पनियों के लिए राज्य को सूचना देना अनिवार्य होता है।

ग. गणनाकारों का प्रशिक्षण

गणनाकारों के लिए प्रशिक्षण-कार्यक्रम होने चाहिए, ताकि वे प्रश्नों को सही ढंग से समझ सकें तथा उत्तरदाताओं को पूछताछ का उद्देश्य तथा अध्ययन का महत्त्व समझ सकें। उन्हें अपने व्यवहार में विनम्र होने का प्रशिक्षण भी मिलना चाहिए।

घ. मार्गदर्शी सर्वेक्षण (Pilot Survey)

यदि बड़े पैमाने पर सर्वेक्षण किया जाना है तब पहले छोटे पैमाने पर एक सर्वेक्षण (जिसे मार्गदर्शी सर्वेक्षण कहते हैं) किया जाना चाहिए।

मार्गदर्शी सर्वेक्षण के गुण

मार्गदर्शी सर्वेक्षण के निम्नलिखित गुण हैं :

i. इसके द्वारा प्रश्नावली में पूछे जाने वाले प्रश्नों की उपयोगिता का पूर्व-परीक्षण किया जा सकता है।
ii. बड़े पैमाने के सर्वेक्षण के दौरान उत्पन्न होने वाली किसी आकस्मिक समस्या से बचा जा सकता है।
iii. इससे गणनाकारों की क्षमता को आँका जा सकता है और यदि उनकी कुछ समस्याएँ हों तो उनका निराकरण किया जा सकता है।
iv. प्रश्नावली के प्रश्नों पर उत्तरदाताओं की प्रतिक्रियाओं को जाना जा सकता है।
v. वास्तविक सर्वेक्षण में लगने वाले समय का अनुमान किया जा सकता है।
vi. बड़े पैमाने के सर्वेक्षण की कुल लागत का अनुमान लगाया जा सकता है।
vii. आँकड़ों के कुछ पक्षों का एक प्रारंभिक अनुमान लगाया जा सकता है।

3. गणना विधि

गणना विधि को 'पूर्ण-गणना विधि' भी कहते हैं। इस विधि की मुख्य विशेषता यह है कि इसके अंतर्गत जनसंख्या की प्रत्येक व्यक्तिगत इकाई का समावेश होता है। ऐसा नहीं होता कि हम कुछ इकाइयों को चुन लें तथा शेष को छोड़ दें।

इसी आधार पर, भारत में जनगणना दस वर्षों में एक बार होती है। घर-घर में पूछताछ की जाती है, जिसमें भारतवर्ष का प्रत्येक परिवार सम्मिलित रहता है। जनांकिकीय आँकड़े जैसे जन्म-दर, मृत्यु-दर, जनसंख्या का आकार तथा संघटन आदि एकत्र किए जाते हैं तथा 'रजिस्ट्रार जनरल' के ऑफिस से प्रकाशित किए जाते हैं। अभी जनगणना फरवरी 2001 में की गई है। समय-समय पर विनिर्माण इकाइयों की गणना भी की जाती है। गणना विधि का प्रयोग भारत में प्रमुख फसलों के अंतर्गत कुल क्षेत्रों का अनुमान लगाने के लिए भी किया जाता है। ये आँकड़े ग्रामीण रिकार्डों से प्राप्त किए जाते हैं जिन्हें गाँव का मुखिया उचित रूप से रखता है।

जनगणना की प्रक्रिया के अंतर्गत पहले अध्ययन के उद्देश्य के अनुसार प्रश्नावली तैयार की जाती है तथा इन प्रश्नावलियों को प्रशिक्षित गणनाकारों की एक बड़ी टीम के साथ क्षेत्रों में उत्तरदाता के पास भरने के लिए भेजा जाता है। इसके लिए एक निश्चित समयावधि होती है।

गणना-विधि द्वारा संग्रहीत आँकड़ों की संख्या अत्यधिक होती है, इसलिए निष्कर्ष निकालने के लिए, संचालन की सुविधा के लिए पहले हमें उन्हें संक्षिप्त करना पड़ता हैं। दूसरे शब्दों में, पहले संभावित त्रुटियों की दृष्टि से आँकड़ों की जाँच की जाती है उसके बाद उन्हें उपयुक्त रूप से सारणीबद्ध किया जाता है तथा यदि आवश्यक हो तो, चित्रों द्वारा प्रस्तुत किया जाता है।

आँकड़े संग्रह करने की प्रक्रिया में इन त्रुटियों की संभावना रहती है —

आँकड़े संग्रहण में त्रुटियाँ

क. मापन संबंधी त्रुटियाँ

एक सरल उदाहरण के रूप में मान लें कि आपकी कक्षा के प्रत्येक विद्यार्थी को शिक्षक के मेज की लंबाई मापनी है। प्रत्येक विद्यार्थी को एक-एक मापक टेप दिया जाए। यदि आप प्रत्येक विद्यार्थी द्वारा लिए गए माप की तुलना करें तो देखेंगे कि सभी माप समान नहीं हैं। मापों में अंतर के कई कारण हो सकते हैं। कुछ विद्यार्थियों ने इकाई के निकटतम तक मापा है, जबकि कुछ ने दशमलव के दसवें स्थान के निकटतम तक। त्रुटियों को पूर्णांकित करने की आवश्यकता होती है। माप में अंतर मापक फीतों में बनावट के दोष के कारण भी हो सकता है। कुछ विद्यार्थी असावधान भी हो सकते हैं।

किसी परिवार के सर्वेक्षण के लिए, मान लें कि उपभोग के विभिन्न मदों पर परिवार के व्यय की जानकारी लेनी है। यदि हम परिवार के मुखिया से खाद्य-सामग्रियों पर मासिक व्यय पूछें तो हमें निकटवर्ती (approximate) आँकड़े ही प्राप्त होंगे। इसी प्रकार, मान लें कि संतरे की कीमतों के आँकड़े एकत्र करने हैं। हम जानते हैं कि कीमतें प्रत्येक दुकान तथा प्रत्येक बाजार में बदलती हैं। कीमतें वस्तुओं के गुणों के अनुसार भी बदलती हैं। प्रश्न यह है कि कौन-सी कीमतें ली जाएँ। अतः इसके लिए हमें औसत कीमतें निकालने की आवश्यकता होती है।

प्रायोगिक विज्ञानों (जैसे भौतिकी, रसायनशास्त्र आदि) में विभिन्न उपकरणों द्वारा माप लेने के कारण ऐसी त्रुटियों की संभावना रहती है।

ख. प्रश्नावली के दुरुपयोग संबंधी त्रुटियाँ

गणनाकार, या, उत्तरदाता प्रश्नावली के कुछ प्रश्नों का गलत अर्थ लगा सकते हैं। चूँकि ऐसे सर्वेक्षणों में मापन बड़े पैमाने पर होता है, अतः गणनाकारों की एक बड़ी टीम की आवश्यकता पड़ती है। सभी गणानाकार समान रूप से कुशल नहीं होते। कुछ असावधान लोग, प्रशिक्षण के बावजूद, अपने कार्य को गंभीरता से नहीं लेते। सामूहिक-कार्य में समन्वयन की कमी भी हो सकती है।

ग. रेकार्डिंग संबंधी त्रुटियाँ

गणनाकार या उत्तरदाता आँकड़े लिखने में गलती कर सकता है। उदाहरण के लिए वह 13 के स्थान पर 31 लिख सकता है आदि। कभी-कभी लिखावट इतनी गंदी और अस्पष्ट होती है कि गणनाकार उत्तरों को फाइल अथवा कंप्यूटर में स्थानांतरित करते समय गलत पढ़ लेता है।

घ. अनुत्तर संबंधी त्रुटियाँ

ये त्रुटियाँ तब पैदा होती हैं, जब उत्तरदाता प्रश्नावली को भरने से इनकार कर देता है अथवा उत्तरदाता गणनाकार के बार-बार जाने पर भी नहीं मिलता है। अनुत्तर की त्रुटियों की संभावना अधिक तब रहती है, जब गणनाकार प्रश्नावली को उत्तरदाता के पास स्वयं न ले जाकर, डाक द्वारा भेजता है। ऐसी स्थिति में उत्तरदाता प्रश्नावली को भली-भाँति भरकर लौटाने पर ध्यान नहीं देता है या आलस्यवश भेजने की परवाह नहीं करता है। कई बार प्रश्नावलियाँ डाक में ही खो जाती हैं।

ङ. अंकगणितीय त्रुटियाँ

यदि कुछ प्रश्नों में गणित के हिसाब-किताब की आवश्यकता हो, तो त्रुटियों की संभावना रहती है। उदाहरण के लिए, यदि प्रश्न है कि 'पिछले महीने खाद्यसामग्रियों पर कितना व्यय हुआ था'? इसके लिए परिवार के मुखिया को चावल, गेहूँ, नमक, चीनी, दूध, फल व सब्जियों पर हुए व्यय को भी जोड़ना पडता है। इन वस्तुओं तथा इनकी कीमतों

को याद करने में तथा उनकी कीमतों को जोड़ने में त्रुटियाँ हो सकती हैं।

पूर्ण गणना (या गणना) विधि द्वारा क्षेत्रीय सर्वेक्षणों के दौरान ऊपर दी गई त्रुटियों की संभावना काफी अधिक रहती है क्योंकि त्रुटियाँ संचयित होती जाती हैं तथा गणनाकारों के बड़े समूह को पर्याप्त प्रशिक्षण देना तथा उनके कार्यों का समन्वयन व निरीक्षण करना कठिन हो सकता है।

गणना-विधि से आँकड़ों को संग्रहीत करने की लागत अधिक होती है क्योंकि गणनाकारों के एक बड़े समूह को प्रशिक्षित करना तथा उनके क्षेत्रीय कार्यों का समन्वयन व निरीक्षण करना पड़ता है। गणनाकारों की यात्रा पर भी काफी खर्च की आवश्यकता होती है। इसके अतिरिक्त खाने-पीने, स्टेशनरी आदि पर होने वाले खर्च भी हैं।

अनेक परिस्थितियों में गणना करना संभव नहीं होता जैसे, जंगल में प्राप्त होने वाली कुल लकड़ियों या नदी में मछलियों या अभयारण्य के पक्षियों की कुल संख्या का अनुमान लगाने में।

इन स्थितियों में हम **प्रतिदर्श आँकड़ों** के आधार पर केवल अनुमान लगा सकते हैं।

4. प्रतिदर्श विधि

मान लें कि हम किसी क्षेत्र के लोगों की औसत आय जानना चाहते हैं।

जनगणना-विधि के अनुसार, उस क्षेत्र के प्रत्येक व्यक्ति की आय की जानकारी लेकर, उन्हें जोड़कर, उस क्षेत्र के व्यक्तियों की संख्या से भाग देना होता है। यह विधि बहुत खर्चीली है क्योंकि गणनाकारों की अधिक संख्या की आवश्यकता होती है। ऊपर दी गई त्रुटियों के द्वारा परिणामों के दूषित होने की भी संभावना रहती है।

अतः इसके स्थान पर **प्रतिदर्श विधि** का प्रयोग भी किया जा सकता है। इस विधि के अनुसार, एक क्षेत्र के कुछ व्यक्तियों का एक प्रतिनिधि प्रतिदर्श चुन कर उनकी आय ज्ञात की जा सकती है। चुने हुए व्यक्तियों के समूह की औसत आय से पूरे क्षेत्र के व्यक्तियों की औसत आय का 'अनुमान' लगाया जा सकता है।

सामान्यतः, प्रतिदर्श विधि में, समष्टि से कुछ व्यक्तियों का एक प्रतिनिधि प्रतिदर्श चुनते हैं तथा प्रतिदर्श आँकड़ों के आधार पर समष्टि की विशेषताओं का 'अनुमान' प्राप्त करते हैं। उदाहरण के लिए, हम पंजाब में गेहूँ की उपज का अनुमान चुने हुए खेतों में गेहूँ की उपज के आधार पर कर सकते हैं; किसी जंगल की कुल लकड़ियों का अनुमान जंगल के कुछ पेड़ों को काटकर कर सकते हैं, आदि।

प्रतिदर्श विधि के गुण

क. पूर्ण-गणना विधि की तुलना में सर्वेक्षण का खर्च बहुत कम होगा ।

ख. आँकड़ों का संग्रह करने, उन्हें सारणीबद्ध करने तथा उनका विश्लेषण करने में काफी कम समय लगेगा।

ग. त्रुटियों की संभावना कम होगी।

घ. गणनाकारों के एक छोटे समूह की आवश्यकता होगी, जिससे उन्हें प्रशिक्षित करना, उनके कार्यों का निरीक्षण करना तथा उन्हें संभावित त्रुटियों से बचाना आसान होगा।

प्रतिचयन-त्रुटियाँ (Sampling errors)

जब समष्टि की विशेषताओं को जानने के लिए, प्रतिदर्श विधि का प्रयोग किया जाता है, तो दूसरे प्रकार की त्रुटियों की संभावना होती है। इन्हें **प्रतिचयन त्रुटियाँ** (Sampling errors) कहते हैं। ऊपर बताई गई मापन त्रुटियों, रेकार्डिंग त्रुटियों आदि को **अप्रतिचयन त्रुटियाँ** (Non-sampling errors) कहते हैं।

उदाहरण के लिए, किसी क्षेत्र के व्यक्तियों की औसत आय का अनुमान जो व्यक्तियों के एक

छोटे समूह की आय के आधार पर निकाला गया हो, उस क्षेत्र के लोगों की वास्तविक औसत आय के बराबर नहीं होती है। किसी क्षेत्र के प्रतिदर्श अनुमान तथा वास्तविक आय के अनुमान के अंतर को प्रतिचयन-त्रुटि कहते हैं।

यह ध्यान में रखना आवश्यक है कि समष्टि की विशेषताओं (औसत आय, आय में परिवर्तन आदि) के अलग-अलग अनुमान प्राप्त होंगे। इसे प्रतिचयन के कारण उतार-चढ़ाव या प्रतिचयन उच्चावचन (sampling fluctuation) कहते हैं।

उदाहरण

एक ऐसी स्थिति पर विचार करें जहाँ समष्टि में केवल पाँच व्यक्ति हों। चर x के मापक 10, 15, 20, 25 तथा 30 हैं। हम देखते हैं कि समष्टि औसत है —

$$\frac{10+15+20+25+30}{5} = \frac{100}{5} = 20$$

मान लें कि हम समष्टि औसत का अनुमान इनमें से दो के प्रतिनिधि प्रतिदर्श के द्वारा निकालना चाहते हैं। समष्टि के 5 व्यक्तियों में से 2 के प्रतिदर्श निकालने के $\frac{5 \times 4}{2} = 10$ तरीके हैं। अतः 10 संभावित प्रतिदर्श निम्नलिखित हैं —

प्रतिदर्श	x के मूल्य	प्रतिदर्श	x के मूल्य
1	10, 15	6	15, 25
2	10, 20	7	15, 30
3	10, 25	8	15, 25
4	10, 30	9	15, 30
5	10, 35	10	15, 30

अलग-अलग प्रतिदर्शों द्वारा प्राप्त समष्टि औसत के अनुमान तथा उनकी प्रतिचयन-त्रुटियाँ निम्नलिखित सारणी में दी गई हैं :

प्रतिदर्श	अनुमान	प्रतिचयन त्रुटि
1	12.5	-7.5
2	15	-5
3	17.5	-2.5
4	20	0
5	17.5	-2.5
6	20	0
7	22.5	2.5
8	22.5	2.5
9	25	5
10	27.5	7.5

अनुमान की प्रतिचयन त्रुटि = अनुमान - वास्तविक औसत, जहाँ वास्तविक औसत = 20

प्रतिदर्श विधि की अनिवार्य शर्त है कि **विचाराधीन विशेषता की दृष्टि से प्रतिदर्श समष्टि का प्रतिनिधि होना चाहिए।** उदाहरण के लिए मान लें हम किसी शहर के लोगों की औसत आय का अनुमान करना चाहते हैं। हम जानते हैं कि किसी शहर में कुछ ऐसे इलाके होते हैं, जहाँ बहुत धनी व्यक्ति रहते हैं तथा कुछ ऐसे, जहाँ केवल बहुत गरीब लोगों की झोपड़ियाँ ही होती हैं। कुछ मध्यवर्गीय लोग भी होते हैं। हमारे प्रतिदर्श द्वारा आय की भिन्नता अवश्य स्पष्ट होनी चाहिए। यदि हमारे प्रतिदर्श में किसी विशेष वर्ग – बहुत धनी या बहुत गरीब – की संख्या अधिक है, तो औसत आय पूर्वग्रह ग्रस्त हो सकता है। इसी प्रकार, पंजाब में गेहूँ की उपज के अनुमान की जानकारी लेते समय हमें ध्यान रखना चाहिए कि सिर्फ बड़े या सिर्फ छोटे खेत ही न चुने जाएँ, आदि।

निम्नलिखित अनुच्छेदों में हम प्रतिनिधि प्रतिचयन निकालने की कुछ विधियों की चर्चा करेंगे।

5. प्रतिनिधि प्रतिचयन निकालने की विधियाँ

क. यादृच्छिक प्रतिचयन (Random Sampling)

सबसे पहले, यह जान लें कि यादृच्छिक प्रतिचयन का अर्थ अव्यवस्थित प्रतिचयन नहीं, जहाँ चुनाव का कोई नियम लागू नहीं होता। वस्तुतः यादृच्छिक प्रतिचयन एक शुद्ध वैज्ञानिक तकनीक पर आधारित है।

यादृच्छिक प्रतिचयन में समष्टि की प्रत्येक व्यक्तिगत इकाई के प्रतिचयन में शामिल होने की समान संभावना होनी चाहिए।

चुनाव प्रक्रिया द्वारा या शर्त पूरी होनी चाहिए। चुनाव की प्रक्रियाओं का वर्णन नीचे किया जा रहा है।

लॉटरी विधि

मान लें कि आपकी कक्षा में 50 विद्यार्थी हैं तथा आपको 5 विद्यार्थियों का एक यादृच्छिक प्रतिचयन निकालना है। इसके लिए आप निम्नलिखित विधि अपना सकते हैं –

क. कागज के एक ही आकार-प्रकार के 50 चिट तैयार करें।

ख. इन चिटों पर विद्यार्थियों के नाम लिखें (एक चिट पर एक विद्यार्थी का नाम)।

ग. इन चिटों को एक बक्से में डालकर अच्छी तरह मिलाएँ।

घ. अंत में, बक्से में बिना देखे, पाँच चिटें निकालें।

ङ. जिन विद्यार्थियों के नाम निकाली गई चिटों पर हों वे ही आवश्यक यादृच्छिक प्रतिचयन का निर्माण करते हैं।

लॉटरियों में लोगों द्वारा क्रय किए गए टिकटों में नंबर दिए रहते हैं। ये टिकट एक बक्से में डालकर मशीन द्वारा अच्छी तरह मिला लिए जाते हैं। उसके बाद एक निश्चित परिमाण में टिकट निकाले जाते हैं। जिन लोगों के पास निकाले गए टिकटों के नंबर के प्रतिपत्रक (counterfoil) होते हैं, वे ही विजयी होते हैं।

लॉटरी विधि के दोष

लॉटरी विधि के निम्नलिखित दोष हैं :

क. यदि समष्टि में इकाइयों की संख्या बहुत अधिक हो, तो लॉटरी विधि का प्रयोग कठिन हो जाता है।

ख. यदि समष्टि का आकार असीमित हो तो लॉटरी विधि का प्रयोग संभव नहीं होता है।

उदाहरण के लिए, यदि किसी अभयारण्य से 50 पक्षियों या किसी नदी से 50 मछलियों या किसी नीम के पेड़ से 50 पत्तियों का प्रतिचयन निकालना हो तो लॉटरी विधि का प्रयोग नहीं किया जा सकता। यदि किसी शहर के 50 परिवारों, या किसी राज्य में किसी विशेष फसल को उपजाने वाले 50 खेतों का प्रतिचयन निकालना हो, तो यह विधि उपयुक्त नहीं।

यादृच्छिक संख्याओं का प्रयोग

कई बार यादृच्छिक प्रतिचयन निकालने के लिए यादृच्छिक संख्याओं का प्रयोग किया जाता है।

यादृच्छिक संख्याएँ विशेष अंकगणितीय विधियों द्वारा बनाई गई हैं ताकि समष्टि में प्रत्येक यक्तिगत इकाई के चुने जाने की बराबर संभावना हो। यादृच्छिक संख्याएँ पुस्तकों के रूप में प्रकाशित हैं तथा उपयुक्त सॉफ्टवेयर पैकेज द्वारा भी प्राप्त की जा सकती हैं। यादृच्छिक संख्या-सारणी की महत्त्वपूर्ण विशेषता यह है कि हम किसी भी पृष्ठ को खोलकर पंक्ति की ओर से या कॉलम की ओर से पढ़ना शुरू कर सकते हैं। इसके द्वारा जो संख्याएँ प्राप्त होती हैं वे यादृच्छिक संख्याएँ हैं। इसके लिए आवश्यकतानुसार 2-अंक, 3-अंक या 4-अंक यादृच्छिक संख्याओं का प्रयोग कर सकते हैं। यादृच्छिक संख्या-सारणी का एक नमूना इस

पुस्तक के परिशिष्ट में दिया गया है।

निम्नलिखित उदाहरण द्वारा यादृच्छिक संख्याओं के प्रयोग को समझाया गया है।

उदाहरण

50 विद्यार्थियों की एक कक्षा से 5 विद्यार्थियों का एक यादृच्छिक प्रतिचयन निकालने के लिए, हम विद्यार्थियों को 01, 02, ..., 50 तक की संख्याएँ प्रदान करते हैं। इसके लिए हम दो अंकों की यादृच्छिक संख्या-सारणी का प्रयोग करते हैं। मान लें कि हम यादृच्छिक संख्या सारणी को पहले पृष्ठ की दसवीं पंक्ति से पढ़ना शुरू करते हैं (देखें परिशिष्ट 'क' सारणी (i) पृष्ठ संख्या 107)। हमें 57, 60, 86, 32 तथा 44 संख्याएँ मिलती हैं। इनमें से कुछ संख्याएँ ऐसी हैं जो किसी भी विद्यार्थी को नहीं मिली हैं। अतः हम एक नियम बनाते हैं कि जब भी प्राप्त संख्या 50 से ऊपर हो, तो विद्यार्थियों की संख्या यादृच्छिक-सारणी से प्राप्त की गई संख्या से 50 को घटाकर निकाली जा सकती है जैसे 57-50 = 07, 60-50 = 10, 86-50 = 36। यदि संख्या 50 से कम हो तो कोई समस्या ही नहीं। दूसरा तरीका यह हो सकता है कि हम 50 से ऊपर की संख्याओं को छोड़ दें तथा आगे की संख्या लें।

01,......50 तक की संख्या देने के स्थान पर विद्यार्थियों की नामांकन संख्या का भी प्रयोग किया जा सकता है। यदि वे 3-अंक की संख्याएँ हैं, तो 3-अंक यादृच्छिक संख्या-सारणी का प्रयोग किया जा सकता है।

उदाहरण

पंजाब में गेहूँ उपजाने वाले 5 खेतों को चुनना है। गाँव के रेकार्ड में सभी खेतों को एक 'खसरा संख्या' दी जाती है। अतः आवश्यकतानुसार 2-अंक, 3-अंक यादृच्छिक संख्या-सारणी का प्रयोग कर सकते हैं।

ख. स्तरित यादृच्छिक प्रतिचयन (Stratified Random Sampling)

स्तरित यादृच्छिक प्रतिचयन के निम्नलिखित चरण हैं –

i. संपूर्ण समष्टि को सजातीय स्तरों में उपविभाजित किया जाता है।
ii. प्रत्येक स्तर से यादृच्छिक प्रतिचयन विधि द्वारा व्यक्तिगत इकाइयों का एक निश्चित अनुपात (जैसे 1% या ½% आदि) निकालते हैं।

इस विधि का प्रयोग तब होता है जब समष्टि विचाराधीन विशेषता के संदर्भ में सजातीय नहीं है। उदाहरण के लिए, किसी शहर के लोगों की औसत आय का अनुमान लगाने के क्रम में हमने देखा कि कुछ स्थानों पर बहुत धनी व्यक्ति रहते हैं तथा कुछ में बहुत गरीब। कुछ स्थानों पर मध्यवर्गीय व्यक्ति रहते हैं। आय-स्तर की दृष्टि से समष्टि सजातीय नहीं है। इसके लिए सजातीय स्तरों की पहचान करनी होगी (उन स्थानों की अलग-अलग पहचान करनी होगी जहाँ बहुत धनी, मध्यवर्गीय तथा बहुत गरीब लोग रहते हैं)। उसके बाद प्रत्येक स्तर से व्यक्तियों का एक निश्चित अनुपात निकालते हैं।

इस विधि के लिए विचाराधीन विशेषता के संदर्भ में प्रत्येक स्तर की आपस में सजातीयता आवश्यक शर्त है, भले ही ये स्तर एक-दूसरे से बहुत भिन्न हों।

अभ्यास

1. 'प्राथमिक' तथा 'द्वितीयक' आँकड़ों में अंतर बताएँ। द्वितीयक आँकड़ों के किन्ही तीन स्रोतों के नाम बताएँ।

2. सांख्यिकीय सारणियाँ, जो जिलावार जन्म तथा मृत्यु-दर (प्रति 1000 जनसंख्या में जन्म तथा मृत्यु की संख्या) दिखाती हैं, भारत की जनगणना 2001 के प्रकाशनों से ली गई हैं। ये प्राथमिक आँकड़े कहे जाएँगे या द्वितीयक ?
3. क्षेत्रीय सर्वेक्षण की योजना बनाने में प्रयोग की जाने वाली मुख्य विधियों के नाम लिखें।
4. गणना विधि से क्षेत्रीय सर्वेक्षण में किन त्रुटियों की संभावना की कल्पना की जा सकती है ?
5. साक्षात्कार विधि तथा डाक प्रश्नावली विधि द्वारा प्राथमिक आँकड़े एकत्र करने के गुणों और दोषों का वर्णन करें।
6. निम्नलिखित की परिभाषाएँ बताएँ : (i) अन्वेषणकर्त्ता, (ii) गणनाकार, (iii) उत्तरदाता।
7. प्रतिचयन तथा अप्रतिचयन त्रुटियों में अंतर बताएँ।
8. मापक त्रुटियाँ तथा रेकार्डिंग त्रुटियाँ क्या हैं?
9. आँकड़े संग्रह करने में त्रुटियों के प्रमुख स्रोत क्या-क्या हैं?
10. गणना विधि की तुलना में प्रतिचयन विधि द्वारा आँकड़े एकत्र करने के क्या-क्या लाभ हैं ?
11. 'यादृच्छिक प्रतिचयन' की परिभाषा बताएँ। यह अव्यवस्थित प्रतिचयन से किस प्रकार भिन्न है ?
12. मान लें कि आपकी कक्षा में 10 विद्यार्थी हैं। उनमें से आप तीन को चुनना चाहते हैं। अतः इसके लिए कितने प्रतिचयन बनाए जा सकते हैं ?
13. अपनी कक्षा के 10 विद्यार्थियों में से 3 को चुनने के लिए आप किस प्रकार लॉटरी विधि का प्रयोग करेंगे, बताएँ।
14. क्या लॉटरी विधि द्वारा हमेशा यादृच्छिक प्रतिचयन ही प्राप्त होता है ? वर्णन करें।
15. यादृच्छिक संख्या सारणियों का प्रयोग करते हुए अपनी कक्षा के 10 विद्यार्थियों में से 3 विद्यार्थियों को चुनने की प्रक्रिया का वर्णन करें।
16. 'यादृच्छिक प्रतिचयन' तथा 'स्तरित यादृच्छिक प्रतिचयन' का वर्णन करते हुए इनमें अंतर स्पष्ट करें। दोनों के उदाहरण दें।

अध्याय - 3

आँकड़ों का संगठन

1. आँकड़ों का वर्गीकरण

वर्गीकरण का तात्पर्य है वस्तुओं को उपयुक्त क्रम से व्यवस्थित करना तथा उन्हें सजातीय समूहों में रखना। उदाहरण के लिए, लाइब्रेरी में पुस्तकों तथा पत्रिकाओं का विषयों के आधार पर; विद्यार्थियों का परीक्षा में प्राप्त अंकों के आधार पर; पौधों तथा जानवरों का उनकी संरचना तथा उद्गम के आधार पर वर्गीकरण किया जाता है।

आँकड़ों को समय, स्थान या दोनों के अनुसार व्यवस्थित करते हैं। उदाहरण के लिए राष्ट्रीय आय, कुल उपभोग, जनसंख्या के आकार आदि से संबद्ध काल-श्रेणी आँकड़े। इसके अतिरिक्त भिन्न-भिन्न समय-बिंदुओं पर विभिन्न राज्यों के साक्षरता-दर के आँकड़े, आदि।

हम एकल चर से संबद्ध आँकड़ों के सामूहीकरण की विधियों की चर्चा विस्तार से नीचे के अनुच्छेदों में करेंगे। इसके पहले कुछ ऐसे पदों को जान लें जिनका प्रयोग दैनिक जीवन में होता है परंतु सांख्यिकी में उनका विशेष अर्थ है।

क. चर एवं गुण (Variable and Attribute)

सामान्य भाषा में, चर पद का तात्पर्य ऐसी विशेषताओं से है जो परिवर्तनशील होती हैं। उदाहरण के लिए, अलग-अलग व्यक्तियों के कद तथा उनके रंग-रूप बदलते रहते हैं, अतः वे व्यक्तियों के परिवर्तनशील चर हैं। इसी प्रकार, लोगों की बुद्धि, वस्तुओं की कीमतें तथा आय भी चर हैं।

सांख्यिकी में 'चर' पद का प्रयोग तभी किया जाता है, जब ये परिवर्तनशील विशेषताएँ संख्याओं में मापी जा सकें। अतः लोगों के कद व वजन चर हैं, क्योंकि वे संख्याओं के द्वारा मापे जा सकते हैं। वस्तुओं की कीमतें, समय तथा स्थान के साथ बदलती हैं, तथा उनका संख्यात्मक मापन संभव है। अतः कीमत एक चर है। इसी प्रकार, व्यक्तियों की आय, उपभोग के विभिन्न मदों पर परिवारों के व्यय, परिवारों के आकार, फर्मों की आगतें व निर्गतें, ये सभी चर हैं।

यद्यपि, व्यक्तियों के रंग-रूप, उनकी बुद्धिमत्ता तथा कला व संगीत के प्रति उनकी अभिरुचि बदलती रहती है, ये उस प्रकार संख्यात्मक रूप में नहीं मापे जा सकते हैं जिस प्रकार लोगों के कद व वजन, या कीमतें व आय। अतः ये सांख्यिकी में 'चर' नहीं कहे जाते हैं। इन्हें 'गुण' कहा जाता है। हम व्यक्तियों का वर्गीकरण उनके गुणों की कोटियों के आधार पर कर सकते हैं। विश्लेषण के उद्देश्य से संख्यात्मक मानों के स्थान पर कोटियों का प्रयोग किया जा सकता है।

ख. संतत तथा असंतत (विविक्त) चर (Continuous and Discrete Variable)

कोई चर तब 'संतत' कहा जाता है तब यह किसी दिए गए परास (range) के अंतर्गत कोई भी मूल्य धारण कर सके। यह मूल्य या तो समाकल (integral) मूल्य (पूर्णांक जैसे 1,2,3, ...) हो सकते हैं, या भिन्नात्मक मूल्य (जैसे $\frac{1}{2},\frac{3}{4},\frac{7}{8}$......) या

$\sqrt{2} = 1.414.......,\sqrt{3} = 1.732.......,\sqrt{7} = 2.645....$

जैसे मूल्य हो सकते हैं (जो पूरी तरह भिन्नात्मक मूल्य भी नहीं हैं)। दूसरे शब्दों में, यह एक दिए गए परास के अंतर्गत कोई भी मूल्य धारण कर सकता है। उदाहरण के लिए, व्यक्तियों के कद व वजन, वस्तुओं की कीमतें, लोगों की आय आदि को संतत चर माना जा सकता है (यद्यपि, व्यवहार में ये मापन सन्निकट मान ही होते हैं, जो दशमलव के एक या दो स्थान तक ही लिए जाते हैं, सही मान एक दिए गए परास में कुछ भी हो सकता है)।

यदि चर कुछ मूल्य विशेष ही धारण करे (जैसे पूर्णांक), तो इसे असंतत या विविक्त चर कहते हैं। उदाहरण के लिए, विभिन्न कक्षाओं या विभिन्न स्कूलों में विद्यार्थियों की संख्या या परिवारों के आकार विविक्त चर हैं (क्योंकि वे केवल समाकल मान ही धारण कर सकते हैं) ।

ग. समष्टि (Population)

सामान्य भाषा में 'समष्टि' पद का अर्थ है किसी क्षेत्र में रहने वाले व्यक्तियों की संख्या। हम लोगों की संख्या गिनकर उस क्षेत्र की समष्टि (जनसंख्या) का आकार जान सकते हैं। इसी प्रकार किसी देश के जंगलों में जानवरों की संख्या या किसी बाग में पौधों की संख्या, आदि भी जान सकते हैं। समष्टि (जनसंख्या) का अर्थ है 'कुल संख्या'।

सांख्यिकी में, **किसी क्षेत्र की सभी व्यक्तिगत इकाइयों के लिए एकल चर या चरों के समुच्चय से संबद्ध आँकड़े उस चर या उन चरों की समष्टि का निर्माण करते हैं**। यदि आँकड़े किसी एक चर से संबद्ध हैं तो मापन के समुच्चय उस चर की **एकविचर समष्टि** (Univariate Population) का निर्माण करते हैं। यदि आँकड़ों का संबंध दो चरों से हो तो यह **द्विविचर समष्टि** (Bivariate Population) तथा यदि आँकड़े चरों के एक पूरे समुच्चय से संबंद्ध हों, तो यह **बहुविचर समष्टि** (Multivariate Population) कहलाता है। उदाहरण के लिए, कीमतों की या आय की **एकविचर समष्टि** हो सकती है। किसी क्षेत्र के सभी व्यक्तियों के कद व वजन की **द्विविचर समष्टि** हो सकती है; या परिवारों के उपभोग के विभिन्न मदों पर व्यय की **बहुविचर समष्टि** हो सकती है।

यहाँ हम एकविचर समष्टि तक ही अपनी चर्चा सीमित रखेंगे तथा (क) **आवृत्ति सरणी**, एवं (ख) **आवृत्ति वितरण** के निर्माण का वर्णन करेंगे।

2. आवृत्ति सरणी (Frequency Array)

यदि चर 'अ' एक असंतत चर है तो हम इसकी आवृत्ति सरणी बनाते हैं जिससे हमें 'अ' के प्रत्येक मूल्य की संगत आवृत्तियाँ प्राप्त होती हैं (यहाँ वर्ग-अंतराल नहीं होते हैं)। इसे निम्नलिखित उदाहरण द्वारा समझें।

उदाहरण

परिवारों के आकार (परिवार के सदस्यों की संख्या) जानने के लिए 100 परिवारों का एक सर्वेक्षण किया गया। सर्वेक्षण के परिणाम सारणी 3.1 में आवृत्ति सरणी के रूप में वर्गीकृत किए गए हैं :

सारणी - 3.1

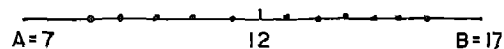
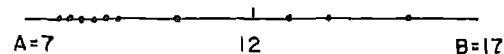
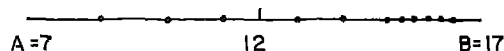
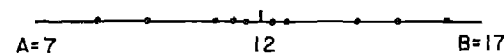
परिवारों के आकार की आवृत्ति सरणी

परिवारों के आकार 'अ'	परिवारों की संख्या
(1)	(2)
1	5
2	15
3	25
4	35
5	10
6	5
7	3
8	2
कुल	100

सारणी 1 के कॉलम (1) में चर 'अ' के मूल्य (परिवारों के आकार) तथा कॉलम (2) में संगत आवृत्तियाँ (परिवारों की संख्याएँ) दी गई हैं। अतः पाँच परिवार ऐसे हैं जिनके सदस्यों की संख्या 1 है; 15 ऐसे, जिनके सदस्यों की संख्या 2 है आदि। सारणी 3.1 परिवारों के आकार की आवृत्ति सरणी दिखाती है।

3. आवृत्ति वितरण (Frequency Distribution)

मान लें कि 'अ' का उच्चतम मूल्य 'ख' तथा निम्नतम मूल्य 'क' है। तब R(परास) = ख-क है जो 'अ' का कुल **परास** है। एक लंबा परास यह दर्शाता है कि 'अ' के मूल्य एक लंबे अंतराल में फैले हुए हैं या 'अ' के मूल्यों का विचरण (Variation) बहुत अधिक है। परास यदि छोटा हो तो इसका अर्थ है कि 'अ' का विचरण कम है। अतः परास 'अ' के विचरण (या परिक्षेपण) का मापक है।

हालाँकि, परास 'अ' के विचरण का एक **स्थूल** मापक है। यह परास के अंतर्गत 'अ' के मूल्यों के वितरण के विषय में कुछ नहीं बताता। क्या 'अ' के मूल्य परास के अंतर्गत समान रूप से वितरित हैं जैसा आरेख 3.1 में दिखाया गया है: या, वे परास की ऊपरी या निचली सीमाओं के निकट या मध्य के निकट के किसी मूल्य के आस-पास एकत्रित हैं, जैसे आरेख 3.2, 3.3 तथा 3.4 में दिखाए गए हैं:

आरेख 3.1 : परास के अंतर्गत समान रूप से वितरित मूल्य

आरेख 3.2 : परास के निम्न भाग में एकत्रित मूल्य

आरेख 3.3 : परास के उच्च भाग में एकत्रित मूल्य

आरेख 3.4 : परास के मध्य में एकत्रित मूल्य

उदाहरण के लिए, मान लें कि हम किसी परीक्षा में एक लाख विद्यार्थियों द्वारा गणित में प्राप्त अंकों पर विचार कर रहे हैं। सबसे अधिक अंक सौ (100) हैं तथा सबसे कम अंक शून्य (0) है। अतः परास (R) = 100 है। यह संभव है कि 70 प्रतिशत विद्यार्थियों को 60 से अधिक, तथा 20 प्रतिशत को 40 से कम अंक मिले हैं। दूसरी स्थिति में, 70 प्रतिशत विद्यार्थियों को 40 तथा 60 के बीच अंक मिले हैं, आदि।

दूसरे उदाहरण में, मान लें कि हमारे पास 10,000 व्यक्तियों की मासिक आय के आँकड़े हैं;

जिसका अधिकतम 50,000 तथा न्यूनतम 1000 है। अतः परास 49,000 है। हम देखते हैं कि अधिकतर व्यक्तियों (लगभग 70 प्रतिशत) की आय 5000 के आस-पास है तथा बहुत कम लोगों (लगभग 2 प्रतिशत) की आय 30,000 के आसपास है।

परास के अंतर्गत अंकों के वितरण की एक बेहतर जानकारी के लिए, हमें कुल परास को वर्ग-अंतरालों में उप-विभाजित कर देना चाहिए तथा विभिन्न वर्गों में मूल्यों की संख्याओं की जानकारी लेनी चाहिए। **किसी वर्ग-विशेष में मूल्यों की संख्या उस वर्ग की आवृत्ति (Frequency) कहलाती है।**

इसके द्वारा आवृत्ति वितरण का निर्माण होता है। आवृत्ति वितरण का निर्माण करते समय निम्नलिखित बिंदुओं पर ध्यान देना चाहिए –

क. वर्ग-अंतरालों की संख्या कितनी हो,
ख. प्रत्येक वर्ग-अंतराल का आकार क्या हो,
ग. वर्ग-सीमाओं का चुनाव कैसे करें।

वर्गों की संख्या

इसके बारे में कोई निश्चित नियम नहीं है, परंतु एक कार्यकारी नियम के तहत वर्गों की संख्या 5 से 15 के बीच होनी चाहिए। ध्यान रहे, कि वर्गों की संख्या बहुत अधिक होगी, यदि हम वर्ग-अंतरालों के आकार छोटे रखें तथा वर्गों की संख्या कम होगी, यदि वर्ग-अंतरालों के आकार बड़े रखें।

उदाहरण के लिए, यदि परास 70 है तथा वर्ग-अंतरालों के आकार 2 रखे जायँ तो हमें $70 \div 2 = 35$ वर्गों की आवश्यकता होगी। यदि वर्ग-अंतरालों के आकार 5 रखे जाएँ तो हमें $70 \div 5 = 14$ वर्गों की आवश्यकता होगी।

यदि आवृत्ति सारणी में 15 से अधिक वर्ग-अंतराल हों तो इसका आकार बहुत बड़ा हो जाएगा, जो ठीक नहीं। यदि वर्ग-अतंरालों की संख्या 5 से कम होती वर्ग-अंतरालों में आँकड़ों के वर्गीकरण का उद्देश्य ही पूरा नहीं होगा।

वर्ग-अंतरालों के आकार

वर्ग-अंतराल समान आकार के हो सकते हैं या समान आकार के न भी हो सकते हैं। यहाँ पर हम केवल समान वर्ग-अंतरालों की चर्चा करेंगे। ऐसी स्थिति में (जहाँ वर्ग-अंतराल समान हों), जैसे ही वर्गों की संख्या निर्धारित हो जाए, वर्ग-अतंराल का आकार निश्चित हो जाता है।

मान लें कि n वर्गों की संख्या है तथा सभी वर्ग-अंतरालों का आकार h है, तो $R = n \times h$। परास R तथा वर्गों की संख्याएँ n जानने के बाद, हम $h = \frac{R}{n}$ को वर्ग-अंतराल का आकार मान सकते हैं। यदि परास 70 है तथा हम 10 वर्ग बनाते हैं, तो वर्ग-अंतराल का आकार 7 होगा।

यह आवश्यक है कि चर के मूल्य वर्ग-अंतरालों के बीच समान रूप से वितरित हों। ऐसी स्थिति में, हम यह मान सकते हैं कि ये मूल्य वर्ग-अंतरालों के मध्य मानों के बराबर हैं। उदाहरण के लिए, यदि वर्ग-अंतराल 10-20 है, तो उस वर्ग के सारे मूल्यों के मध्य मान (Mid Values) 15 (जो $\frac{20+10}{2} = 15$ के द्वारा प्राप्त होता है) के बराबर होंगे। ऐसा मानने से त्रुटियों की संभावना कम होगी, क्योंकि धनात्मक त्रुटियाँ ऋणात्मक त्रुटियों को बराबर कर देंगी। परंतु, यदि ये मूल्य वर्ग-अंतरालों में समान रूप से वितरित न हों तो त्रुटियाँ अधिक होंगी।

वर्ग-अंतरालों का चुनाव

मान लें कि x एक ऐसा **संतत** चर है जो किसी दिए गए परास में कोई भी मूल्य धारण कर सकता है। ऐसी स्थिति में, वर्ग-अंतरालों की ऐसी सीमाओं का

चुनाव संभव है जो किसी भी प्रेक्षित मूल्य के बराबर नहीं है। उदाहरण के लिए, व्यक्तियों के कद एक संतत चर है, यद्यपि, व्यवहार में, लंबाई को इकाई मूल्य के निकटतम तक (से.मी.) मापा जा सकता है जैसे 165, 170, 169, 171, ..., या दशमलव के दसवें स्थान के निकटतम तक जैसे 165.3, 170.4, 168.9, 170.8, ...। हम वर्ग-अंतरालों को इस प्रकार दिखा सकते हैं – 160.55-165.55, 165.55-175.55, ..., ताकि x का कोई भी प्रेक्षित मूल्य किसी भी वर्ग-अंतराल के बराबर न हो।

चूँकि x एक **असंतत** चर है जो केवल समाकल मान ही धारण करता है, अतः समाकल मूल्यों के अतिरिक्त अन्य वर्ग-सीमाओं के चुनाव का कोई अर्थ नहीं। उदाहरण के लिए, मान लें कि परिवार का आकार x है तो किसी सर्वेक्षण में x का मूल्य 5, 4, 8, 3, 7, ... हो सकता है। इस स्थिति में हम निम्नलिखित विधियों में से किसी एक का चुनाव कर सकते हैं :

समावेशित विधि (Inclusive Method)

हम वर्ग-अंतरालों का चुनाव इस प्रकार करते हैं कि ऊपरी तथा निचली सीमाएँ अंतराल का अंग बन जाएँ। दूसरे शब्दों में, **वर्ग-सीमाएँ भी समावेशित (अंतर्विष्ट) हों**। उदाहरण के लिए, मान लें कि हम वर्ग-अंतराल 1-5, 6-10, 11-15, 16-20 आदि रखते हैं; तथा यह भी देखते हैं कि एक ऐसा परिवार है जिसका आकार (सदस्यों की संख्या) 5 है। यह परिवार पहले वर्ग अंतराल 1-5 में सम्मिलित होगा क्योंकि दोनों ही सीमाएँ सम्मिलित हैं। इसी प्रकार, वह परिवार, जिसका आकार 6 है, दूसरे वर्ग-अंतराल 6-10 में सम्मिलित है।

अपवर्जित विधि (Exclusive Method)

इस विधि में वर्ग-अंतरालों की ऊपरी या निचली सीमा में से एक सम्मिलित नहीं होती है। मान लें वर्ग-अंतराल 1-5, 5-9, 9-13, 13-17 आदि हैं तथा हम स्पष्ट कर देते हैं कि वर्ग अंतराल में ऊपरी सीमा सम्मिलित नहीं है तथा निचली सीमा सम्मिलित है। अतः एक परिवार जिसका आकार 5 हो, दूसरें वर्ग-अंतराल 5-9 में सम्मिलित होगा।

इसी प्रकार हम यह भी स्पष्ट कर सकते हैं कि ऊपरी सीमा सम्मिलित है तथा निचली सीमा सम्मिलित नहीं है। ऐसी स्थिति में, वह परिवार जिसका आकार 5 है, पहले वर्ग (1-5) में संम्मिलित होगा।

4. आवृत्ति सरणी का निर्माण तथा आवृत्ति वितरण

उदाहरण 1 : आवृत्ति सरणी

किसी वर्ग के बीस विद्यार्थियों ने शिक्षक का मूल्यांकन करने का निश्चय किया। मूल्यांकन 1 से 5 के स्केल पर किया गया, जहाँ '1' का अर्थ 'सबसे अधिक अच्छा' तथा '5' का अर्थ ' सबसे कम अच्छा' है। परिणाम आवृत्ति-सारणी में नीचे दिए गए हैं।

कोटियाँ (x)	1	2	3	4	5
विद्यार्थियों की संख्या (f)	4	6	7	2	1

उपर्युक्त स्थिति में, चर (x) कोटियाँ दर्शाता है। यह एक असंतत चर है, जिसके मूल्य हैं 1,2,3,4 तथा 5 । विद्यार्थियों की संख्या आवृत्तियाँ दिखाती हैं। 4 विद्यार्थियों ने शिक्षक को '1' अंक दिया है तथा 6 ने '2' अंक आदि। कुल आवृत्तियाँ 20 हैं।

यदि हम मिलान चिह्नों का प्रयोग करें तो आवृत्तियों की गिनती आसान हो जाएगी। इसे हम सारणी 3.2 आवृत्ति-सरणी द्वारा स्पष्ट कर सकते हैं :

सारणी - 3.2

कोटियों की आवृत्ति सरणी

कोटियाँ x	मिलान चिह्न	आवृत्ति f
1	////	4
2	~~////~~ /	6
3	~~////~~ //	7
4	//	2
5	/	1
कुल		20

यहाँ प्रत्येक विद्यार्थी के लिए एक मिलान-चिन्ह का प्रयोग किया गया है। 4 विद्यार्थी ऐसे हैं जिन्होनें शिक्षक को पहली श्रेणी में रखा है, अतः x=1 के सामने 4 मिलान चिह्न लगाए गए हैं; 6 विद्यार्थी ऐसे हैं जिन्होंने शिक्षक को दूसरी श्रेणी में रखा है, अतः x=2 के समाने 6 मिलान चिह्न लगाए गए हैं; आदि। सुविधा के लिए पाँचवाँ मिलान चिह्न, पहले चारों मिलान चिह्नों को काटते हुए लगाया जाता है। इससे गिनती में आसानी होती है।

उदाहरण 2 : आवृत्ति वितरण

सारणी 3.3 में किसी परीक्षा में 100 विद्यार्थियों द्वारा गणित में प्राप्त प्रतिशत अंक दिए गए हैं।

सारणी 3.3 में दिए गए आँकड़ों के लिए एक आवृत्ति-वितरण का निर्माण करें।

जैसा पहले बताया जा चुका है, वर्गों की संख्या, उनके आकार तथा वर्ग-अंतराल चुनने के कई तरीके हैं।

मान लें कि हम 10 वर्गों का चुनाव करते हैं, जैसे 0-10, 10-20, ..., 90-100 । सभी वर्ग-अंतरालों के आकार 10 हैं, तथा वर्ग-अंतरालों की ऊपरी वर्ग-सीमाएँ अगले वर्ग-अंतराल की निचली सीमाओं के बराबर है। अतः सुविधा के लिए हम एक नियम बनाएँ कि **वर्ग-अंतरालों की ऊपरी सीमाएँ सम्मिलित नहीं हैं परंतु निचली सीमाएँ सम्मिलित हैं**। अतः यदि एक विद्यार्थी ऊपरी सीमा के बराबर अंक प्राप्त करता है, तो हम उसे अगले वर्ग में सम्मिलित करते हैं।

आवृत्तियों की गिनती विभिन्न वर्ग-अंतरालों के सामने मिलान-चिह्न लगाकर की जाती है। मिलान-चिह्न (/) वर्ग के सामने हर उस विद्यार्थी

सारणी - 3.3

किसी परीक्षा में 100 विद्यार्थियों द्वारा गणित में प्राप्त अंक

47	45	10	60	51	56	66	96	49	40
60	59	56	55	62	48	59	55	51	41
42	69	64	66	50	59	57	65	62	50
64	30	37	75	17	56	20	14	55	90
62	51	55	14	25	34	90	49	56	54
70	47	49	82	40	82	60	85	65	66
49	44	64	69	70	48	12	28	55	65
49	40	25	41	71	80	09	56	14	22
66	53	46	70	43	61	59	12	30	35
45	44	57	76	82	39	32	14	90	25

के लिए लगाया जाता है जिसके अंक उस वर्ग-अंतराल में हों। उदाहरण के लिए, यदि किसी विद्यार्थी के अंक 57 हों, तो वर्ग-अंतराल 50-60 के सामने एक मिलान-चिह्न लगाया जाता है, यदि अंक 71 हैं तो वर्ग-अंतराल 70-80 के सामने मिलान-चिह्न लगाया जाता है। यदि किसी को 40 अंक मिले हों तो वर्ग 40-50 के सामने मिलान-चिह्न लगाया जाता है।

गिनती की सुविधा के लिए, चार मिलान-चिह्नों को इस प्रकार लगाया जाता है (////) तथा पाँचवें को इन चारों को काटते हुए (////)। अतः यदि किसी वर्ग में 16 मिलान-चिह्न हों तो हम उन्हें इस प्रकार लगाते हैं (//// //// //// /)।

किसी वर्ग में आवृत्तियाँ उस वर्ग के सामने लगे मिलान-चिह्नों के बराबर होंगी। सारणी 3.4 में आवृत्ति वितरण दिखाया गया है।

सारणी - 3.4

किसी परीक्षा में 100 विद्यार्थियों द्वारा गणित में प्राप्त प्रतिशत अंकों का आवृत्ति वितरण

अंकों का प्रतिशत (x) वर्ग-अंतराल	मिलान-चिह्न	आवृत्ति (f)
0-10	/	1
10-20	//// ///	8
20-30	//// /	6
30-40	//// //	7
40-50	//// //// //// //// /	21
50-60	//// //// //// //// ///	23
60-70	//// //// //// ////	19
70-80	//// /	6
80-90	////	5
90-100	////	4
कुल		100

उदाहरण 3 : आवृत्ति वितरण

40 व्यक्तियों की दैनिक मजदूरी आय के आँकड़े नीचे दिए गए हैं :

200	120	350	550	400	140	350	85
180	110	110	600	350	500	450	200
170	90	170	800	190	700	630	170
210	185	250	120	180	350	110	250
430	140	300	400	200	400	210	300

हम 'दैनिक मजदूरी आय' को एक संतत चर मान सकते हैं। हम देखते हैं कि अधिकतम मजदूरी 880 रु. है तथा न्यूनतम 85 रु. । अतः परास 715 है। वर्ग अंतरालों के आकार h=100 मानकर, वर्ग-अंतराल निम्नलिखित विधि से बना सकते हैं – 70.5-170.5, 170.5-270.5, ... 770.5-870.5। सारणी 3.5 में आवृत्ति वितरण दिखाया गया है।

सारणी - 3.5

दैनिक मजदूरी आय का आवृत्ति वितरण

मजदूरी (रु. में) (x)	मिलान-चिह्न	आवृत्ति (f)
70.5-170.5	//// //// //	12
170.5-270.5	//// //// /	11
270.5-370.5	//// /	6
370.5-470.5	////	5
470.5-570.5	//	2
570.5-670.5	//	2
670.5-770.5	/	1
770.5-870.5	/	1
कुल		40

5. सामूहीकरण की त्रुटि (Error of Grouping)

आँकड़ों का वर्गीकरण वर्ग-अंतरालों में करने पर आवृत्ति वितरण प्राप्त होता है। आवृत्ति वितरण विभिन्न वर्ग-अंतरालों में प्रेक्षणों (आवृत्तियों) की संख्या दिखाता है, **उनका वास्तविक मूल्य नहीं।** किसी भी वर्ग-अंतराल में सभी मूल्य वर्ग-अंतराल के मध्य मूल्यों के बराबर मान लिए जाते हैं। इसे **सामूहीकरण-त्रुटि** कहते हैं।

उदाहरण के लिए, सारणी 3.5 में, वर्ग 70.5 से 170.5 में 12 मूल्य लिए गए हैं, जैसे:

120	140	85	110	110	170
90	170	170	120	110	140

हम इन सब को वर्ग-अंतराल के मध्य मान के बराबर मान लेते हैं: $\frac{1}{2}(70.5 + 170.5) = 120.5$ अतः प्रत्येक वर्ग में, सामूहीकरण त्रुटियाँ निम्नलिखित होंगी :

-0.5	19.5	-35.5	-10.5	-10.5	49.5
-30.5	49.5	49.5	-0.5	-10.5	19.5

सारणी 3.6 में, विभिन्न वर्ग-अंतरालों में मजदूरी आय के वास्तविक आँकड़े दिए गए हैं :

सारणी - 3.6

विभिन्न वर्ग-अंतरालों में वास्तविक मजदूरी आय

मजदूरी (X)	वास्तविक मजदूरी आय	आवृत्ति
70.5-170.5	120, 140, 85, 110, 110, 170, 90, 170, 170, 120, 110, 140	12
170.5-270.5	200, 180, 200, 190, 210, 185, 250, 180, 250, 200, 210	11
270.5-370.5	350, 350, 350, 350, 300, 300	6
370.5-470.5	400, 450, 430, 400, 400	5
470.5-570.5	550, 500	2
570.5-670.5	600, 630	2
670.5-770.5	700	1
770.5-870.5	800	1
कुल		**40**

यहाँ ध्यान देना चाहिए कि किसी भी वर्ग अंतराल में सामूहीकरण की त्रुटियाँ कम होंगी, यदि सारे मूल्य वर्ग-अतंरालों के अंतर्गत समान रूप से वितरित हों तथा वर्ग-अंतराल बहुत लंबे न हों। ऐसी स्थिति में धनात्मक तथा ऋणात्मक त्रुटियाँ एक-दूसरे को निष्क्रिय कर देंगी। इसे ध्यान में रखते हुए, विद्यार्थी सारणी 5 तथा सारणी 6 के सामूहीकरणों की जाँच कर सकते हैं तथा यदि इन्हें संतोषजनक न पाएँ तो इसके बेहतर वर्गीकरण प्रस्तुत कर सकते हैं। इसके लिए क्या विद्यार्थी समान वर्ग-अंतरालों का चुनाव करेंगे या असमान वर्ग-अंतरालों का?

आवृत्ति-सारणी में सामूहीकरण-त्रुटि का प्रश्न नहीं उठता क्योंकि वहाँ वर्ग-अंतराल होते ही नहीं।

अभ्यास

1. 'आँकड़ों के वर्गीकरण' से आप क्या समझते हैं?
2. चर तथा गुण के बीच अंतर बताएँ। उदाहरण द्वारा स्पष्ट करें।
3. संतत तथा असंतत चरों के बीच अंतर बताएँ। उदाहरण द्वारा स्पष्ट करें।
4. सांख्यिकी में प्रयुक्त होने वाले 'समष्टि' पद का अर्थ स्पष्ट करें। एकविचर, द्विविचर तथा बहुविचर समष्टि की परिभाषाएँ बताएँ। इनके उदाहरण दें।
5. आवृत्ति वितरण से आप क्या समझते हैं? आवृत्ति वितरण के निर्माण के संबंध में कौन-कौन-सी मुख्य बातों का ध्यान रखा जाना चाहिए?
6. सामान्यतः हमें कितने वर्ग-अंतरालों का चुनाव करना चाहिए? वर्ग-अंतरालों के आकार के संबंध में किन बातों के आधार पर निर्णय लिया जाना चाहिए?

7. आवृत्ति-सरणी तथा आवृत्ति वितरण में अंतर बताएँ।
8. आँकड़ों के वर्गीकरण में प्रयुक्त होने वाली 'समावेशित' तथा 'अपवर्जित' विधियों का वर्णन करें।
9. किसी विश्वविद्यालय की इंजीनियरिंग प्रवेश परीक्षा में उत्तीर्ण होने वाले 50 विद्यार्थियों के अंकों के प्रतिशत निम्नलिखित हैं:

95	92	91	90	88
82	86	87	87	88
70	72	76	77	77
79	79	79	78	78
68	67	67	66	65
63	63	63	63	64
58	59	60	60	60
60	61	62	62	62
55	55	55	56	57
54	53	52	52	50

अ. प्रतिशत अंकों के वितरण का परास ज्ञात करें।

ब. आँकड़ों को निम्नलिखित वर्ग-अंतरालों में वर्गीकृत करें :

क. 45-55, 55-65, 65-75, ...

ख. 50-55, 55-60, 60-65, ...

ग. प्रत्येक स्थिति में सामूहीकरण की त्रुटि की जाँच करें।

10. निम्नलिखित आँकड़े 50 परिवारों के मासिक पारिवारिक व्यय (रु. में) से संबद्ध हैं

904	1559	3473	735	2760
2041	612	753	1855	4439
5090	1085	823	2346	1523
211	1360	1110	2152	1183
1218	1315	1105	628	2712
4248	812	264	1183	1171
1007	1180	953	137	2048
2025	1583	1324	9621	9676
1397	832	962	2177	2575
1293	365	1146	5222	1396

क. मासिक पारिवारिक व्यय के वितरण का परास ज्ञात करें।

ख. परास को उपयुक्त वर्ग-अंतरालों में बाँट कर व्यय का आवृत्ति वितरण ज्ञात करें।

ग. उपर्युक्त आँकड़ों के लिए 'समान' या 'असमान' कौन-सा वर्ग-अंतराल उपयुक्त होगा ? अपने उत्तर के लिए कारण बताएँ।

घ. उन परिवारों की संख्या ज्ञात करें जिनका मासिक व्यय —

i. 1500 रु. से कम है।

ii. 6000 रु. से अधिक है।

11. इस पुस्तक का कोई भी पृष्ठ निकालें। उपयुक्त वर्ग-अंतरालों को चुनकर, विभिन्न शब्दों में प्रयुक्त अक्षरों की संख्या का आवृत्ति वितरण ज्ञात करें। संख्याओं की गिनती यहाँ नहीं करनी है।

आँकड़ों का प्रस्तुतीकरण

1. विवरणात्मक रूप

यदि प्रेक्षणों की संख्या कम हो तो हम उनका वर्णन सीधे शब्दों में कर सकते हैं। उदाहरण के लिए, 25 अक्तूबर 2001 के टाइम्स ऑफ इंडिया के तीसरे पृष्ठ की रिपोर्ट के अनुसार "राष्ट्रीय संक्रामक रोग संस्थान में एन्थ्रेक्स की आशंका के शिकार 12 रोगी भर्ती हुए........जिससे कुल रोगियों की संख्या बढ़कर 93 हो गई। इन 12 रोगियों में से 6 दिल्ली के, 4 गाजियाबाद के तथा 2 बुलंदशहर के थे। अब तक 52 रोगियों की जाँच हो चुकी है।...., सभी की रिपोर्टें नकारात्मक हैं।" इसी समाचारपत्र के तीसरे पृष्ठ पर 26 अक्तूबर 2001 को प्रकाशित अन्य रिपोर्ट के अनुसार, "ट्रैफिक पुलिस ने इस वर्ष 15 अक्तूबर तक 16891 ऑटो-रिक्शा चालकों को अधिक भाड़ा वसूलने के अपराध में चालान किया। यह संख्या पिछले वर्ष की संख्या 1877 से नौ गुनी अधिक है। सवारियों को ले जाने से मना करने के जुर्म में चालान किए गए ऑटो-रिक्शा चालकों की संख्या में भी तीन गुनी वृद्धि हुई है। सवारियों को ले जाने से मना करने के जुर्म में 15 अक्तूबर तक 19960 ड्राइवरों को सजा दी गई। पिछले वर्ष सिर्फ 5805 ड्राइवरों को सजा दी गई थी।"

सामान्यतः गणना (census) या प्रतिचयन (sample) विधि से एकत्र किए गए आँकड़े जटिल हो जाते हैं। यदि इनके माध्यम से समष्टि की विशेषताओं का अनुमान लगाना है तो इन्हें संक्षिप्त एवं सुव्यवस्थित रूप से प्रस्तुत करने की आवश्यकता है।

इसके लिए आँकड़ों को **सारणीबद्ध** या **आरेखी** रूप में प्रस्तुत किया जा सकता है।

2. सारणीबद्ध प्रस्तुतीकरण

एक संक्षिप्त टिप्पणी के साथ पंक्तियों तथा स्तंभों में आँकड़ों के प्रस्तुतीकरण द्वारा किसी सारणी का निर्माण होता है। एक सरल उदाहरण के तौर पर किसी शहर के 255 परिवारों के लिए एकत्र किए गए आँकड़े निम्नलिखित विधि से सारणीबद्ध किए गए हैं :

सारणी - 4.1

परिवारों के आकारों का वितरण

परिवारों के आकार	2	3	4	5	6	7	8	9	कुल
परिवारों की संख्या	5	25	40	65	50	35	20	15	255

इस सारणी से हम निम्नलिखित निष्कर्ष निकाल सकते हैं :

i. सबसे बड़े परिवार का आकार 9 है तथा सबसे छोटे का 2, अतः परिवारों के आकार 2 से 9 के बीच में हैं;

ii. तीन-चौथाई परिवारों (75%) के आकार 3 से अधिक तथा 8 से कम हैं; तथा

iii. अधिकांश परिवारों (65) में 5 सदस्य हैं।

आँकड़ों को बिना सारणीबद्ध किए, ऐसे सरल निष्कर्ष निकालना भी आसान नहीं होता।

क. सारणी के अंग

सारणी के निम्नलिखित अंग होते हैं :

i. सारणी संख्या (Table Number)

सारणी संख्या या तो सारणी के ऊपर या शीर्षक के पहले दी जाती है। इसका प्रयोग सारणी की पहचान करने के लिए किया जाता है तथा यह संख्याओं के रूप में होती है।

सामान्यतः सारणी-संख्याएँ पूर्णांकों के रूप में दी जाती हैं। हालाँकि कभी-कभी पादांकों जैसे 2.1 तथा 3.1 आदि का भी प्रयोग किया जाता है। ऐसी स्थिति में, पहला अंक उस अध्याय या अनुच्छेद को बताता है जिसमें सारणी दी गई है तथा दूसरा अंक सारणी की क्रम-संख्या को। उदाहरण के लिए, सारणी 2.1 का अर्थ है, दूसरे अध्याय (या अनुच्छेद) की पहली सारणी। इसी प्रकार, सारणी 3.2 का अर्थ है, तीसरे अध्याय (या अनुच्छेद) की दूसरी सारणी।

ii. शीर्षक (Title)

शीर्षक संक्षेप में सारणी की विषय-वस्तु बताता है। अतः यह स्पष्ट शब्दों में होना चाहिए तथा इस प्रकार बनाया जाना चाहिए कि सारणी की महत्त्वपूर्ण विशेषताओं का वर्णन कर सके। यह सारणी के ऊपर सारणी-संख्या के साथ या उसके ठीक नीचे दिया जाना चाहिए। उदाहरण के लिए सारणी 4.1 को देखें।

iii. स्तंभ शीर्षक (Column Heading)

स्तंभ-शीर्षक छोटे उपशीर्षक के रूप में एक संक्षिप्त विवरण देते हैं, जैसे 'जनसंख्या', 'सकल राष्ट्रीय उत्पाद' आदि।

किसी स्तंभ में आँकड़ों के मापन की इकाइयाँ या तो उपशीर्षकों के साथ या उनके नीचे कोष्ठकों में दी जाती हैं। यदि विभिन्न स्तंभों में दी गई सभी संख्याओं के मापन की इकाइयाँ समान हों तो इन्हें प्रत्येक स्तंभ में अलग-अलग देने की आवश्यकता नहीं। ऐसी स्थिति में हम इन्हें सारणी के शीर्षक के नीचे एक ओर लिख सकते हैं। स्तंभों की संख्याएँ, स्तंभों के शीर्षकों के नीचे दी जाती हैं।

iv. पंक्ति शीर्षक (Row Heading)

सारणी की प्रत्येक पंक्ति का एक शीर्षक होता है। उदाहरण के लिए, सारणी 4.2 में, प्रत्येक पंक्ति में सबसे बायीं ओर दिखाए गए 'राज्य' पंक्ति-शीर्षक हैं।

पंक्ति-शीर्षक का एक संक्षिप्त विवरण (जैसे सारणी 4.2 में 'प्रमुख भारतीय राज्य') सारणी के बायें हाथ के ऊपर कोने में दिया गया है।

v. सारणी का काय या क्षेत्र (Body of the Table)

यह सारणी का मुख्य भाग है जिसमें आँकड़े दिए होते हैं। हम किसी भी संख्या की पहचान उस पंक्ति तथा स्तंभ के द्वारा कर सकते हैं जिसमें यह स्थित होता है। उदाहरण के लिए, सारणी 4.2 के अनुसार, आंध्रप्रदेश में 'पुरुषों' की साक्षरता दर 55% तथा पश्चिम बंगाल में 'महिलाओं' की साक्षरता दर 47% हैं।

vi. पाद टिप्पणियाँ (Foot Notes)

पाद टिप्पणियाँ आँकड़ों की कुछ विशेषताओं का वर्णन करने के लिए सारणियों के नीचे दी जाती हैं। यदि ये सारणी की किसी विशेष संख्या से संबंध रखती हों तो उन संख्याओं के ऊपर एक (*) या संख्या लिख दी जाती है। कभी-कभी पाद-टिप्पणी का उद्देश्य सारणी की कुछ महत्त्वपूर्ण विशेषताओं का उल्लेख करना भी होता है।

vii. स्रोत टिप्पणी (Source Note)

यह सारणी में आँकड़ों का स्रोत बताने के लिए होता है। यह सारणी के नीचे दिया जाता है।

सारणी - 4.2

साक्षरता दर

प्रमुख भारतीय राज्य*	साक्षरता दर (%)#		
	कुल	पुरुष	महिला
(1)	(2)	(3)	(4)
आंध्र प्रदेश	44	55	3
असम	53	62	43
बिहार	39	53	2
गुजरात	61	73	49
हरियाणा	56	69	41
कर्नाटक	56	67	44
केरल	90	94	86
मध्य प्रदेश	44	58	29
महाराष्ट्र	65	77	52
उड़ीसा	49	63	35
पंजाब	59	66	50
राजस्थान	39	55	20
तमिलनाडु	63	74	51
उत्तर प्रदेश	42	56	25
पश्चिम बंगाल	58	68	47
भारत	**52**	**64**	**39**

*प्रमुख राज्य वे हैं जिनकी जनसंख्या **1991 की जनगणना** के अनुसार एक करोड़ या अधिक है।

7 वर्ष या उससे अधिक आयु की जनसंख्या के लिए, जम्मू व कश्मीर को छोड़कर।

स्रोतः भारत की जनगणना 1991

ख. सारणी का निर्माण (Construction of a table)

सारणी सामान्यतः स्वतः स्पष्ट होनी चाहिए। किसी सारणी का निर्माण करते समय निम्नलिखित बातों का ध्यान रखना चाहिए :

i. **सारणी सुसंबद्ध, संक्षिप्त तथा एक नजर में पढ़ने योग्य होनी चाहिए।** यह भारी-भरकम और बोझिल नहीं होनी चाहिए। अन्यथा, यह आँकड़ों को संचालनीय रूप में रखने के बुनियादी उद्देश्य को ही पूरा नहीं कर सकेगी।

ii. **सारणी ऐसी हो कि इसके द्वारा आँकड़ों की तुलना संभव हो सके।** जिन संख्याओं की तुलना करनी हो उन्हें आसपास की पंक्तियों व स्तंभों में रखा जाना चाहिए।

iii. **जिन संख्याओं पर जोर देना है उन्हें मोटे अक्षरों में लिखा जाना चाहिए, या उनके चारों ओर एक वृत्त बना देना चाहिए।**

iv. **यदि संख्याएँ बहुत बड़ी हों तो उनका सन्निकटन किया जाना चाहिए (निकटतम करोड़ या लाख तक)**, क्योंकि बड़ी संख्याओं को पढ़ना और उनकी तुलना करना कठिन होता है।

v. **यदि कोई प्रविष्टि 'शून्य' हो तो इसे स्पष्ट किया जाना चाहिए ताकि यह न समझा जाय कि वहाँ पर कोई प्रविष्टि है ही नहीं।**

3. आरेखी प्रस्तुतीकरण (Diagrammatic Form)

आरेख सारणी की तरह सही तो नहीं हो सकते, परंतु ये इसलिए उपयोगी होते हैं क्योंकि इनके द्वारा तुलना करना आसान होता है। इनके द्वारा-

क. तुलनात्मक अध्ययन में शीघ्रता से आँकड़ों के आकार की जानकारी हो जाती है।

ख. आँकड़ों की विशेषताओं की जानकारी हो जाती है।

आरेख के निर्माण में निम्नलिखित बातों का ध्यान रखा जाना चाहिए —

i. **आरेख का एक शीर्षक तथा एक संख्या होनी चाहिए**। ये सामान्यतः या तो आरेख के ऊपर या इसके नीचे दिए जाते हैं। आरेख की संख्या, इसकी पहचान तथा संदर्भ के लिए दी जाती है; तथा शीर्षक संक्षेप में आरेख की विषय-वस्तु बताने के लिए।

ii. **x तथा y दोनों अक्षों पर स्पष्ट लेबल लगाए जाने चाहिए**। आरेख के द्वारा यह स्पष्ट होना चाहिए कि दोनों अक्षों पर क्या मापा जा रहा है। उदाहरण के लिए, हम y-अक्ष पर 'जनसंख्या' या 'सकल राष्ट्रीय उत्पाद' को तथा 'वर्ष' को x-अक्ष पर माप सकते हैं।

iii. **चरों के साथ चरों के मापन की इकाइयाँ भी दी जानी चाहिए**। उदाहरण के लिए, यदि जनसंख्या y-अक्ष पर मापी जा रही है तो हमें y-अक्ष पर 'जनसंख्या (लाख या करोड़ लोग)' तथा x-अक्ष पर 'वर्ष' लिखना चाहिए।

iv. **दोनों अक्षों पर मापन का स्केल आरेख के ऊपर दायें हाथ के कोने पर या आरेख के नीचे दिया जाना चाहिए।** उदाहरण के लिए, 1 से. मी. = 1 लाख या 1 करोड़ जनसंख्या, आदि।

v. **आरेख के द्वारा उद्गम (मूल बिंदु) स्पष्ट किया जाना चाहिए।** यदि बहुत बड़ी संख्याएँ दी हों (जैसे सकल राष्ट्रीय उत्पाद करोड़ रुपयों में या जनसंख्या लाख या करोड़ में, आदि) तो मूल बिंदु 50,000 या 5 लाख रखा जा सकता है। इसे देखकर तुलना करना आसान हो जाता है।

सामान्यतः कई प्रकार के आरेखों का प्रयोग किया जाता है। इनमें से कुछ नीचे दिए जा रहे हैं:

i. **ज्यामितीय आरेख (Geometric Forms)**
   - क. दंड आरेख
   - ख. बहु दंड-आरेख
   - ग. वृत्त आरेख

ii. **आवृत्ति आरेख (Frequency Diagrams)**
   - क. आयत चित्र
   - ख. आवृत्ति बहुभुज
   - ग. आवृत्ति वक्र
   - घ. तोरण

iii. **अंकगणितीय लाइन-ग्राफ (काल श्रेणी आरेख) Arithmetic Line Graphs (Time series Graphs)**

4. ज्यामितीय आरेख

क. दंड आरेख (Bar Diagrams)

दंड आरेख समय, स्थान या समूह के लिए बनाए गए उदग्र (vertical) या क्षैतिज (horizontal) दंड आरेखों के एक समूह को कहते हैं। दंडों की ऊँचाई (या लंबाई) आँकड़ों के आकार दर्शाती है। दंडों की चौड़ाई कुछ भी हो सकती है, परंतु सभी दंडों की चौड़ाई बराबर होनी चाहिए तथा दंड बराबर दूरी पर होने चाहिए।

दंड आरेख विभिन्न समय, स्थान या समूह के बीच तुलना में सहायक होते हैं।

उदाहरण के लिए, आरेख 4.1 में, हम भारत के 15 राज्यों में दंड आरेख द्वारा साक्षरता दर (प्रतिशत में) दिखाने के लिए सारणी 4.2 के कॉलम 2 के आँकड़ों का प्रयोग करते हैं। राज्यों को x-अक्ष पर तथा साक्षरता दर को (1 से.मी. =10%) y-अक्ष पर दिखाया गया है। प्रत्येक राज्य के लिए वास्तविक आँकड़े संगत दण्डों के शीर्ष पर दिए गए हैं।

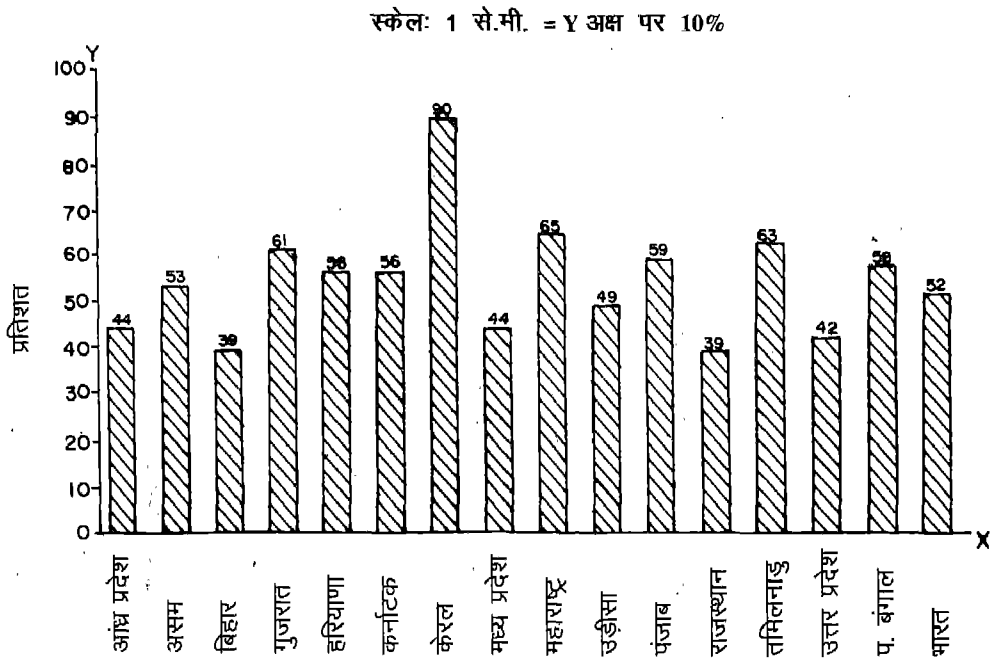
आरेख 4.1 स्पष्ट रूप से दिखाता है कि केरल में साक्षरता-दर अन्य राज्यों से अधिक है। महाराष्ट्र तथा तमिलनाडु का स्थान दूसरा है, यद्यपि इन दोनों राज्यों में साक्षरता दरें काफी कम हैं। बिहार तथा राजस्थान में सारक्षरता दरें निम्नतम हैं।

ख. बहु दंड आरेख (Multiple Bar Diagram)

ये आरेख एक से अधिक चरों के लिए दंड दिखाते हैं। ये एक ही समय में विभिन्न चरों की तुलना करने में सहायक होते हैं। चरों में अंतर दिखाने के लिए हम भिन्न-भिन्न चरों को दर्शाने वाले दंडों में अलग-अलग रंग भरते हैं अथवा उन्हें अलग-अलग प्रकार से छायांकित करते हैं।

बहु दंड आरेख में, दंडों पर अच्छी तरह लेबल लगाए जाने चाहिए या आरेख के एक कोने में दंडों के बारे में अलग से संकेत-सूचकों द्वारा निर्देश दिए जाने चाहिए। उदाहरण के लिए, आरेख 4.2 में सारणी 4.3 में दिए गए आँकड़ों के आधार पर भारत के आयात व निर्यात दिखाए गए हैं।

x-अक्ष पर वर्ष तथा y-अक्ष पर मूल्य (1 से. मी. = 20,000 करोड़ रुपये) दिए गए हैं।

सारणी 4.3 में दंडों की ऊँचाई निर्यात/आयात के मूल्य दिखाती है। प्रत्येक दंड के शीर्ष पर वास्तविक संख्याएँ दी गई हैं।

आरेख 4.1 : राज्य साक्षरता-दर

सारणी - 4.3

भारत में आयात व निर्यात

वर्ष	वर्तमान कीमतों पर आयात (हजार करोड़ रु. में)	वर्तमान कीमतों पर निर्यात (हजार करोड़ रु. में)
(1)	(2)	(3)
1996-97	139	119
1997-98	154	130
1998-99	176(p)	142(p)
1999-00	149*	119*

*अप्रैल-दिसंबर 1999

(p) अनंतिम

स्रोतः *आर्थिक सर्वेक्षण 1999-2000*

एक दूसरे उदाहरण में, सारणी 4.4 में कृषि उत्पादन के सूचकांकों के लिए आरेख 4.3 में एक बहु-दंड आरेख दिखाया गया है।

सारणी - 4.4

कृषि उत्पादन के सूचकांक

(आधार 1981-82 =100 तथा इकाई मूल्य के निकटतम आँकड़े)

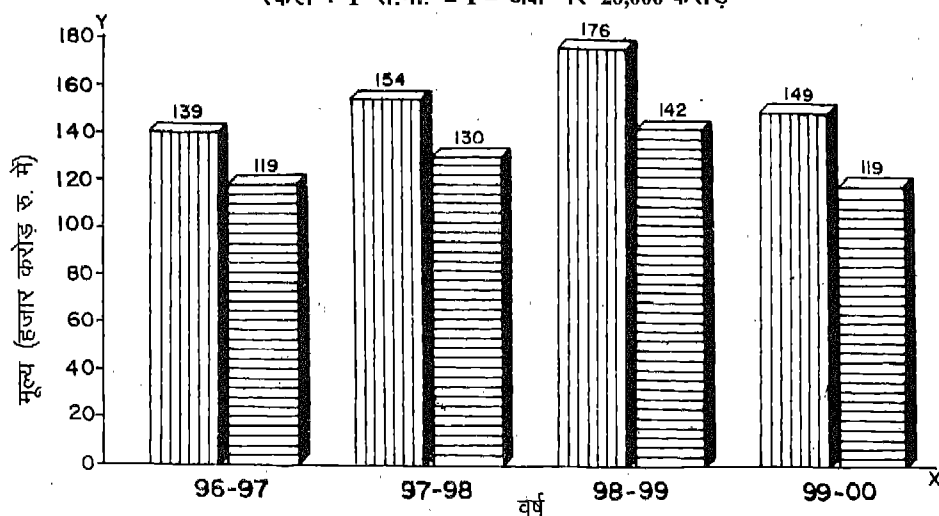
वर्ष	खाद्यान्न		
	चावल	गेहूँ	दालें
(1)	(2)	(3)	(4)
1990-91	149	157	141
1993-94	162	170	131
1994-95	165	187	139
1995-96	155	176	121
1996-97	164	197	140
1997-98	166	189	127
1998-99	173	201	145

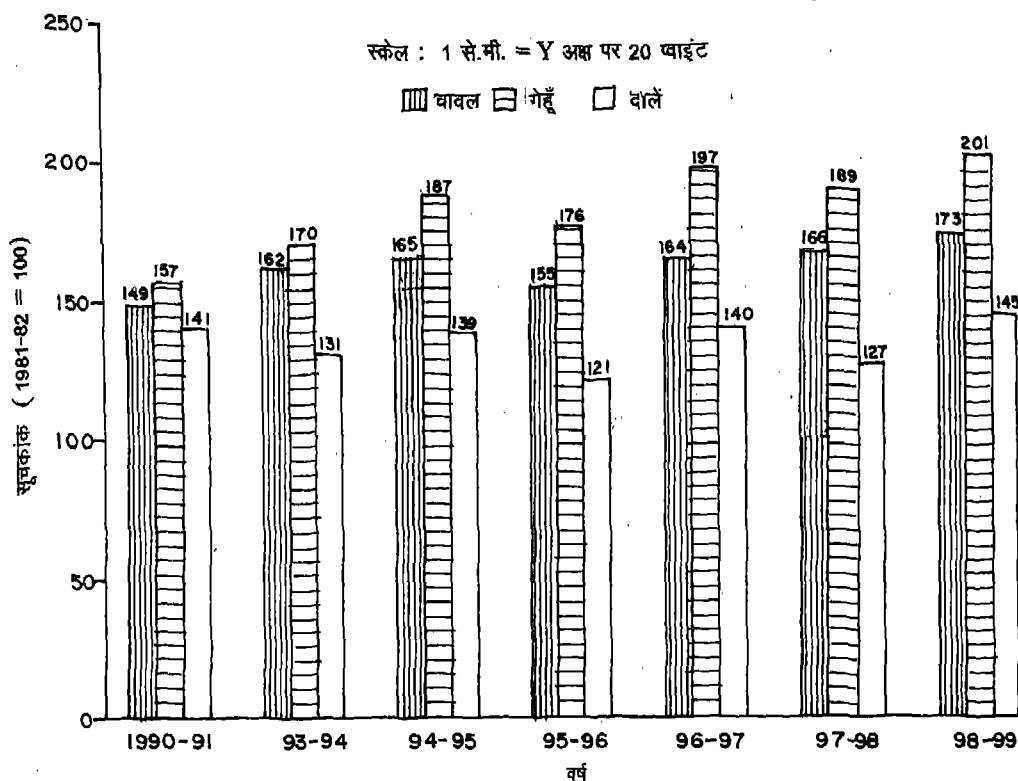
आरेख 4.3 में, वर्ष x-अक्ष पर दिखाए गए हैं। प्रत्येक वर्ष के लिए तीन उदग्र दंड (चावल, गेहूँ तथा दालों के लिए) भिन्न-भिन्न रंगों में बनाए गए हैं। प्रत्येक दंड की ऊँचाई उस दंड के ऊपर दिखाया गया सूचकांक बताती है। प्रत्येक दंड की चौड़ाई समान है। y-अक्ष पर मापन का स्केल 1 से. मी. =20 प्वाइंट। (सुविधा के लिए, सूचकांक उनके इकाई मूल्य के निकटतम तक दिए गए हैं)।

स्केल : 1 से.मी. = Y – अक्ष पर 20,000 करोड़

मूल्य (हजार करोड़ रु. में)

96-97 97-98 98-99 99-00

वर्ष

आयात निर्यात

आरेख 4.2 : आरेख भारत में आयात व निर्यात

आरेख 4.3 : कृषि उत्पादों के सूचकांक

खंडित दंड आरेख (Broken Bar Diagram)

कभी-कभी चरों के परास इतने बड़े होते हैं कि उच्चतम मूल्य के संगत दंड को ग्राफ-पत्र पर खींचना संभव नहीं होता। ऐसी स्थिति में, हम दंडों को कई भागों में बनाते हैं, प्रत्येक नया भाग संख्यात्मक स्केल पर कुछ दूर खाली छोड़कर शुरू होता है। दंड पर लेबल वास्तविक संख्या के अनुसार ही लगाते हैं।

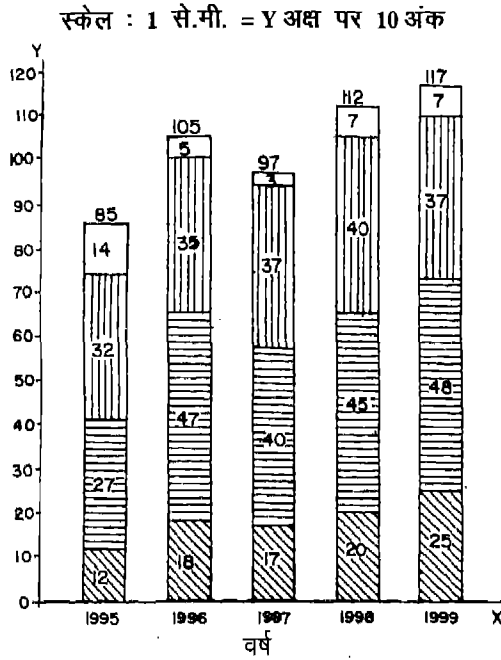
अभ्यास के लिए, 'योजना-व्यय' के लिए प्रथम पंचवर्षीय योजना से आरंभ करके एक दंड-आरेख बनाया जा सकता है। पंचवर्षीय योजनाओं के योजना-व्यय तथा योजनेतर व्यय के लिए द्वि (या बहु) दंड आरेख भी बनाएँ।

घटक दंड आरेख (Component Bar Diagram)

कभी-कभी दंड एक ही चर के विभिन्न घटकों में खण्डित होते हैं। प्रत्येक घटक, पहचान के लिए, भिन्न-भिन्न रंगों/छायाओं/डिजाइनों द्वारा दिखाए जाते हैं।

उदाहरण के लिए, किसी परीक्षा में विद्यार्थियों का परीक्षाफल निम्नलिखित सारणी में दिखाया गया है :

वर्ष	विद्यार्थियों की संख्या	प्रथम श्रेणी	द्वितीय श्रेणी	तृतीय श्रेणी	अनुत्तीर्ण
1995	85	12	27	32	14
1996	105	18	47	35	5
1997	97	17	40	37	3
1998	112	20	45	40	7
1999	117	25	48	37	7

घटक दंड आरेख, नीचे आरेख 4.4 में दिखाया गया है।

आरेख 4.4 : घटक दंड–आरेख

ग. वृत्त आरेख (Pie Diagram)

इसे वृत्त-चार्ट भी कहते हैं। वृत्त-आरेख का प्रयोग विभिन्न वर्गों में कुछ समुच्चयों के प्रतिशत वितरण को बताने के लिए किया जाता है। एक वृत्त किसी दिए हुए प्रतिशत के अनुसार कई खंडों में विभाजित होता है। केन्द्र से संपूर्ण कोण को विशेष अनुपातों में बाँटकर ऐसा किया जाता है। केंद्र पर कुल कोण 360° या 2π होता है, अतः यह चित्र 'वृत्त आरेख' कहलाता है।

उदाहरण के लिए, हम भारत में 1998-99 में कुछ वस्तु-समूहों के आयातों के प्रतिशत वितरण को लेते हैं। आँकड़े सारणी 4.5 में दिए गए हैं तथा वृत्त आरेख आरेख 4.5 में दिखाया गया है। चूँकि समूह 'क' में प्रतिशत 6.1 है तथा केंद्र से कक्षांतरित कोण (angle subtended) $\frac{6.1 \times 360}{100} = 21.96°$ के बराबर है, जो लगभग 22° है। इसी प्रकार हम अन्य वस्तु-समूहों के लिए भी केंद्र से कक्षांतरित कोण निकाल सकते हैं।

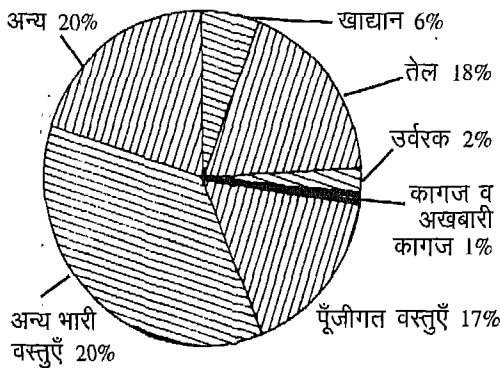
सारणी - 4.5

कुछ वस्तु समूहों के लिए भारत में आयातों का प्रतिशत वितरण

	वस्तु-समूह	प्रतिशत	केंद्र से कोण
1	2	3	4
क.	खाद्यान्न व सहबद्ध उत्पाद	6.1	22°
ख.	तेल	17.6	63°
ग.	उर्वरक	2.3	8°
घ.	विनिर्मित पेपर बोर्ड व अखबारी कागज	1.1	4°
ङ.	पूँजीगत वस्तुएँ	16.6	60°
च.	अन्य भारी वस्तुएँ	35.9	129°
छ.	अन्य	20.4	74°
	कुल	100.00	360°

आयातों का मूल्य = अमेरिकी $ 4,18,579 लाख

स्रोतः *आर्थिक सर्वेक्षण, 1999 – 2000*

एक दूसरा उदाहरण लें। वर्ष 1998–99 में भारत में कुछ वस्तु समूहों के लिए निर्यातों के प्रतिशत विवरण को देखें। आँकड़े सारणी 4.6 में दिए गए हैं तथा आरेख 4.6 में वृत्त-आरेख द्वारा दिखाया गया है।

नोटः ये आयातों के कुल मूल्य के प्रतिशत हैं।

आरेख 4.5 : भारत के प्रमुख आयातों का प्रतिशत वितरण (1998-99)

आरेख 4.6 : आरेख भारत के प्रमुख निर्यातों का प्रतिशत वितरण (1998-99)

सारणी - 4.6

भारत में कुछ वस्तु -समूहों के लिए निर्यातों का प्रतिशत वितरण (1998-99)

	वस्तु-समूह	प्रतिशत	केंद्र से कोण
1	2	3	4
क.	कृषि व सहबद्ध उत्पाद	17.3	62°
ख.	अयस्क व खनिज	2.6	9°
ग.	विनिर्मित वस्तुएँ	77.8	280°
घ.	कच्चा तेल व पेट्रोलियम उत्पाद	0.3	1°
ङ	अन्य	2.0	8°

निर्यातों का कुल मूल्य = अमेरिकी $ 3,36,585 लाख

स्रोतः *आर्थिक सर्वेक्षण, 1999 – 2000*

5. अंकगणितीय लाइन ग्राफ (काल श्रेणी ग्राफ) Arithmetic Line Graph (Time Series Graph)

काल श्रेणी ग्राफ में हम समय की इकाई (वर्ष, माह आदि) x-अक्ष पर तथा चरों के मूल्य y-अक्ष पर मापते हैं।

सारणी 4.7, वर्ष 1986–87 से 1998–99 तक के लिए थोक कीमतों के सूचकांक (आधार वर्ष 1981–82 =100) दिखाती है। काल श्रेणी ग्राफ, आरेख 4.7 में दिखाया गया है। ग्राफ से थोक कीमतों की बढ़ती हुई प्रवृत्ति स्पष्ट है।

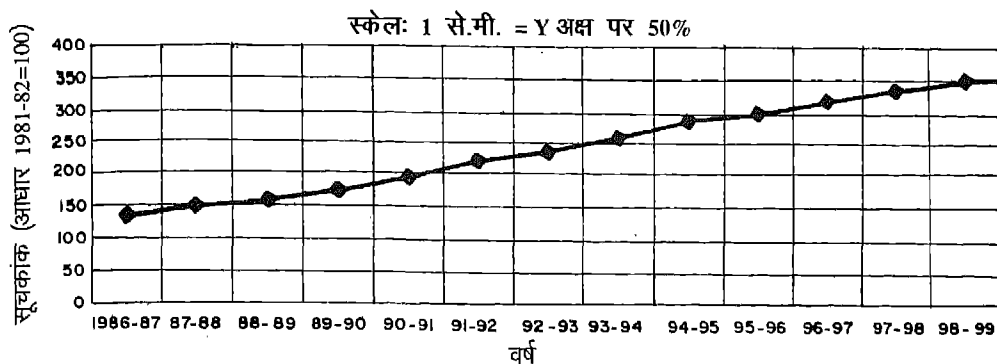
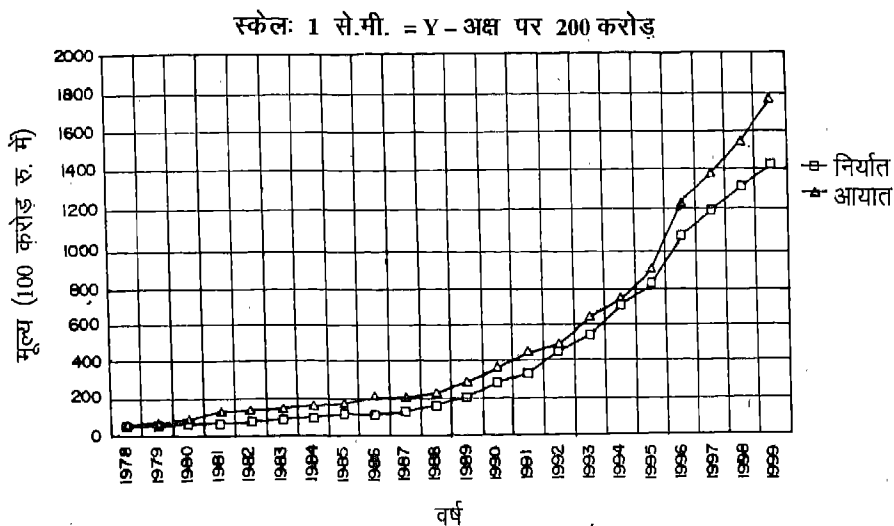
एक दूसरे उदाहरण में वर्ष 1977–78 से 1998–99 तक भारत में निर्यात व आयात के मूल्यों (100 करोड़ रु. में) के लिए सारणी 4.8 के आँकड़े लें। आरेख 4.8 में निर्यात व आयात दोनों के लिए काल श्रेणी ग्राफ दिए गए हैं। हम इन दोनों को अलग-अलग रंगों के प्रयोग द्वारा दिखा सकते हैं या एक को संतत (या गहरी) तथा दूसरी को खंडित लाइनों द्वारा।

सारणी - 4.7

थोक कीमतों के सूचकांक

(आधार 1981–82 = 100)

निम्नलिखित वर्षों का अंतिम सप्ताह	सूचकांक
1986–87	134
1987–88	149
1988–89	157
1989–90	171
1990–91	192
1991–92	218
1992–93	233
1993–94	258
1994–95	285
1995–96	300
1996–97	320
1997–98	337
1998–99	353

स्रोतः *आर्थिक सर्वेक्षण, 1999 –2000*

आरेख 4.7 : थोक कीमतों के सूचकांक

आरेख 4.8 : भारत के निर्यात एवं आयात

सारणी - 4.8

भारत के निर्यातों व आयातों के मूल्य (हजार करोड़ रुपयों में)

वर्ष	निर्यात	आयात
1977-78	54	60
1978-79	57	68
1979-80	64	91
1980-81	67	125
1981-82	78	136
1982-83	88	143
1983-84	98	158
1984-85	117	171
1985-86	109	197
1986-87	125	201
1987-88	157	222
1988-89	202	282
1989-90	277	353
1990-91	326	432
1991-92	440	479
1992-93	532	634
1993-94	698	731
1994-95	827	900
1995-96	1064	1227
1996-97	1186	1369
1997-98	1301	1542
1998-99	1416	1761

अभ्यास

1. निम्नलिखित सूचनाओं को सारणियों के रूप में दिखाएँ :

"सरकारी आँकड़ों के अनुसार पिछले वर्ष (2000) की तुलना में, इस वर्ष (2001) 15 दिसंबर तक (शहर में) अपराधों की संख्या (हत्याओं को छोड़कर) बढ़ गई है।"

"इस वर्ष डकैती के छियासठ (66) केस दर्ज किए गए जबकि पिछले वर्ष इस समय 19 केस दर्ज किए गए थे। हत्या के प्रयासों की संख्या 200 से बढ़कर 256 हो गई, जबकि डकैती की संख्या जो पिछले वर्ष 324 थी, बढ़कर 636 हो गई। चोरी, कारों की चोरी तथा साइकिलों की चोरी की संख्याएँ पिछले वर्ष क्रमशः 2527, 1965, 5129 थीं, बढ़कर क्रमशः 3283, 2759, 5889 हो गई।"

इस वर्ष कुल 51809 केस दर्ज किए गए जबकि पिछले वर्ष केवल 40246 केस दर्ज किए गए थे। आँकड़ों के प्रस्तुतीकरण की 'वर्णनात्मक' तथा 'सारणीबद्ध' विधियों की तुलना करें।

2. प्रश्न 1 में, मान लें कि आप इस अवधि में हुई हत्याओं तथा डकैतियों में हुई वृद्धि को दिखाना चाहते हैं। आप इसे सारणी के रूप में कैसे दिखाएँगे?

3. भारतीय चीनी मिल एसोसिएशन ने रिपोर्ट किया है कि "दिसंबर 2001 के पहले पखवाड़े में चीनी का उत्पादन 387000 टन था, जबकि पिछले वर्ष (2000) के इसी पखवाड़े में यह 378000 टन था........"

"दिसंबर 2001 के पहले पखवाड़े में फैक्टरियों से आंतरिक उपभोग के लिए 283000 टन चीनी खरीदी गई थी तथा निर्यात के लिए 41000 टन, जबकि पिछले वर्ष में इसी पखवाड़े में घरेलू उपभोग के लिए 154000 टन चीनी खरीदी गई थी तथा निर्यात के लिए चीनी नहीं खरीदी गई थी।"

   क. आँकड़ों को सारणीबद्ध करें।
   ख. मान लें कि आपको इन आँकड़ों से आरेख बनाना हो तो आप किस आरेख का चुनाव करेंगे और क्यों?
   ग. इन आँकड़ों का आरेखी प्रस्तुतीकरण करें।

4. क. एक अच्छी सारणी के अपेक्षित गुण क्या हैं?
   ख. "आँकड़ों के प्रस्तुतीकरण के लिए आरेख सारणियों की तरह सही तो नहीं होते पर उनसे अधिक प्रभावशाली होते हैं।" व्याख्या करें।

5. निम्नलिखित को दर्शाने के लिए किस प्रकार के आरेख अधिक प्रभावशाली होते हैं:
   क. किसी वर्ष में मासिक वर्षा
   ख. धर्मों के अनुसार दिल्ली की जनसंख्या
   ग. किसी फैक्टरी में लागत के घटक

6. निम्नलिखित सारणी में भारत में खाद्यान्नों के उत्पादन के निरपेक्ष मूल्य (लाख टन में) दिए गए हैं

वर्ष	खाद्यान्नों का उत्पादन (लाख टन में)
1996-97	1994
1997-98	1923
1998-99	2030
1999-00	1091

आँकड़ों को एक उपयुक्त दंड आरेख द्वारा दिखाएँ।

7. निम्नलिखित सारणी में भारत में विद्धुत उत्पादन (बिलियन KWH) के आँकड़े दिए गए हैं :

वर्ष	विद्युत उत्पादन (बिलियन KWH)
1996–97	394.5
1997–98	420.6
1998–99	448.4
अप्रैल-दिस. 99	355.3

आँकड़ों को एक उपयुक्त दंड आरेख द्वारा प्रदर्शित करें।

8. निम्नलिखित सारणी में कारक लागत पर सकल घरेलू उत्पाद में अनुमानित क्षेत्रकवार वास्तविक वृद्धि दर (पिछले वर्ष की तुलना में प्रतिशत परिवर्तन) दी गई है :

वर्ष	कृषि व सहबद्ध क्षेत्रक	उद्योग	सेवाएँ
(1)	(2)	(3)	(4)
1994–95	5.0	9.2	7.0
1995–96	-0.9	11.8	10.3
1996–97	9.6	6.0	7.1
1997–98	-1.9	5.9	9.0
1998–99	7.2	4.0	8.3
1999–00	0.8	6.9	8.2

मापन का स्केल बताते हुए, आँकड़ों को बहु काल-श्रेणी ग्राफ के रूप में दिखाएँ।

9. निम्न सारणी में 1998-99 में भारत में निर्यातों (प्रतिशत) की दिशा दिखाई गई है :

	गंतव्य स्थान	प्रतिशत
1	2	3
क.	संयुक्त राष्ट्र अमेरिका	21.8
ख.	जर्मनी	5.6
ग.	अन्य यूरोपियन यूनियम सदस्य	14.7
घ.	इंगलैंड	5.7
ङ.	जापान	4.9

1	2	3
च.	रूस	2.1
छ.	अन्य पूर्वी यूरोप	0.6
ज.	तेल उत्पादक राष्ट्र (ओपेक)	10.5
झ.	एशिया	19.0
ञ.	अन्य अल्प विकसित देश	5.6
ट.	अन्य	9.5

कुल निर्यात अमेरिकी $ 33658.5 मिलियन

विभिन्न स्रोतों से निर्यातों के प्रतिशत वितरण को एक वृत्त आरेख द्वारा दिखाएँ।

10. भारत में वर्ष 1998-99 में आयातों के स्रोत (प्रतिशत रूप में) निम्न सारणी में दिए गए हैं :

	स्रोत	प्रतिशत
1	2	3
क.	संयुक्त राष्ट्र अमेरिका	8.7
ख.	जर्मनी	5.1
ग.	अन्य यूरोपियन यूनियन सदस्य	12.4
घ.	इंगलैंड	6.1
ङ.	जापान	5.7
च.	रूस	1.3
छ.	अन्य पूर्वी यूरोप	0.4
ज.	तेल उत्पादक राष्ट्र (ओपेक)	18.7
झ.	एशिया	15.7
ञ.	अन्य अल्प विकसित देश	5.6
ट.	अन्य	20.3

अन्य आयात : अमेरिकी $ 41857 मिलियन

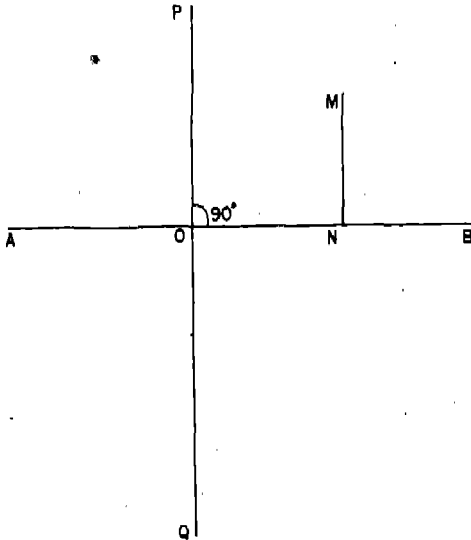
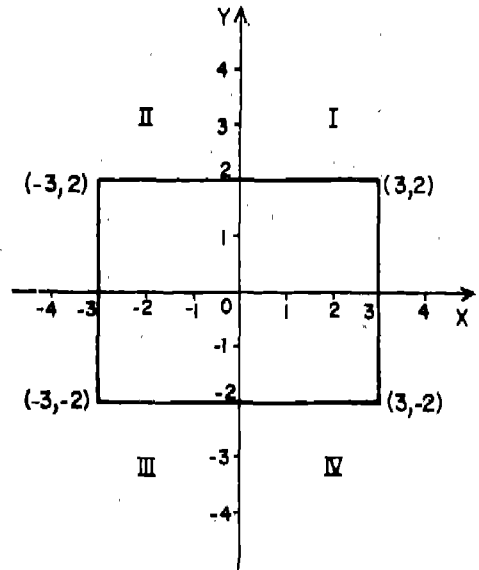
विभिन्न स्रोतों से आयातों के प्रतिशत वितरण को दिखाने के लिए एक वृत्त आरेख बनाएँ।

अध्याय - 5

आवृत्ति वक्र तथा आरेख

1. आँकड़ों का आरेखी प्रस्तुतीकरण

(Diagrammatic Presentation of Data)

किसी वक्र या आरेख को बनाने लिए ग्राफ-पेपर का प्रयोग सुविधाजनक होता है। ग्राफ-पेपर पर किसी बिंदु की स्थिति को दो ऐसी सरल रेखाओं के संदर्भ में बताया जा सकता है जो एक-दूसरे को समकोण द्वारा काटती हैं। परंपरानुसार, उनमें से एक क्षैतिज तथा दूसरा उदग्र होता है। इन संदर्भ रेखाओं को 'अक्ष' तथा इनके 'प्रतिच्छेद' को मूल बिंदु कहते हैं।

जैसा आरेख-5.1 में दिखाया गया है, AB तथा PQ दो अक्ष हैं तथा O मूल बिंदु है। मान लें कि M ग्राफ-पेपर पर कोई बिंदु है। इसकी स्थिति निर्धारित करने के लिए, AB अक्ष पर M से एक लंब MN डालें। ON (मूल बिंदु O से N की दूरी) को M का भुज (abscissa) तथा MN को कोटि (ordinate) कहते हैं। अतः भुज को क्षैतिज-अक्ष AB पर तथा कोटि को उदग्र-अक्ष PQ पर मापते हैं। परंपरानुसार, भुज को x द्वारा तथा कोटि को y द्वारा दिखाया जाता है, अतः AB को x-अक्ष तथा PQ को y-अक्ष कहते हैं।

आरेख 5.1 : निर्देशांक अक्ष

आरेख 5.2 : चार चतुर्थांश

भुज और कोटि दोनों एक साथ किसी बिंदु के निर्देशांक (coordinates) बताते हैं तथा AB, PQ निर्देशांक-अक्ष हैं। पहले किसी बिंदु का भुज बताया जाता है और उसके बाद अक्ष। अतः यदि (2, 4) किसी बिंदु के निर्देशांक हों तो x=2 को भुज तथा y=4 को कोटि कहते हैं।

इस संदर्भ में दूसरी बात यह है कि x-अक्ष का कुछ भाग मूल बिंदु 'o' के दाहिनी ओर तथा कुछ बायीं ओर हैं। 'o' के दाहिनी ओर के भाग x के धनात्मक मूल्य तथा बायीं ओर के भाग ऋणात्मक मूल्य सूचित करते हैं। इसी प्रकार y-अक्ष का कुछ भाग 'o' के ऊपर और कुछ 'o' के नीचे है। y-अक्ष का वह भाग जो 'o' के ऊपर है, y के धनात्मक मूल्य तथा वह भाग जो 'o' के नीचे है, y के ऋणात्मक मूल्य सूचित करता है। आरेख 5.2 को देखें। ग्राफ-पेपर दोनों अक्षों द्वारा चार चतुर्थांशों (quadrants) में विभाजित होता है। चतुर्थांश I के सभी बिंदुओं के निर्देशांक धनात्मक होते हैं (भुज तथा कोटि)। चतुर्थांश II में भुज ऋणात्मक तथा कोटि धनात्मक होता है, चतुर्थांश III में भुज तथा कोटि दोनों ही ऋणात्मक तथा चतुर्थांश IV में भुज धनात्मक तथा कोटि ऋणात्मक होता है।

अभ्यास : ग्राफ-पेपर पर उन बिंदुओं को अंकित करें जिनके निर्देशांक यहाँ दिए गए हैं :
(i) (2, 8), (ii) (8, 2), (iii) (-2, 8), (iv) (-2, -8), (v) (0, 0), (vi) (-4, -6), (vii) (-6, 0), (viii) (0,-6)

2. आवृत्ति सरणी का आरेखी प्रस्तुतीकरण – रेखा चित्र (Diagrammatic Presentation of Frequency Array – Line Graph)

अध्याय 3 में हमने देखा कि यदि चर x असंतत हों, तो हमें एक आवृत्ति सरणी मिलती है जिससे हम x के प्रत्येक मूल्य की आवृत्ति जान सकते हैं। उदाहरण के लिए, किसी शहर में परिवारों के आकार जानने के लिए 100 परिवारों का एक सर्वेक्षण किया गया। सर्वेक्षण के परिणाम निम्नलिखित आवृत्ति सरणी में सारणीबद्ध किए गए हैं :

परिवारों के आकार (x)	परिवारों की संख्या (f)
1	5
2	15
3	25
4	35
5	10
6	5
7	3
8	2
कुल	**100**

आरेख 5.3 : आवृत्ति सरणी का आरेखी प्रस्तुतीकरण

आरेख 5.3 में आवृत्ति सारणी दिखाई गई है। इस आरेख में x के भिन्न-भिन्न मूल्यों पर कोटियों के समुच्चय (परिवारों के आकार) दिए गए हैं, जहाँ कोटि की ऊँचाई संगत आवृत्ति (परिवारों की संख्या) के बराबर है। अतः कोटि की ऊँचाई x=1 पर 5 के बराबर है, x=2 पर 15 के, इत्यादि। हमने y-अक्ष पर माप का स्केल 1 से.मी. =5 परिवार चुना है तथा x -अक्ष पर 1 से.मी. =1 परिवार।

अभ्यास :

मान लें कि किसी घनाकार पासे के पहलू 1, 2, ...6 हैं। इस पासे को 50 बार उछाला गया जिससे निम्नलिखित परिणाम निकले –

x	आवृत्ति
1	6
2	9
3	7
4	10
5	12
6	6
कुल	**50**

आवृत्ति सारणी को आरेख द्वारा दिखाएँ।

3. आयत चित्र (Histogram)

किसी आवृत्ति वितरण को **आयत चित्र** द्वारा दिखाया जा सकता है। इसे **आवृत्ति आयत-चित्र** कहते हैं। इसके लिए हम दो प्रकार की स्थितियों पर विचार करेगें।

स्थिति 1 : जब वर्ग अंतराल समान हों।

स्थिति 2 : जब वर्ग-अंतराल समान नहीं हों।

पहले हम स्थिति-1 को देखेंगे।

उदाहरण 1 : समान वर्ग-अंतराल (Equal Class Intervals)

64 विद्यार्थियों द्वारा गणित में प्राप्त किए गए अंक निम्नलिखित आवृत्ति-वितरण द्वारा दिखाए गए हैं :

सारणी - 5.1

गणित के अंकों का आवृत्ति वितरण

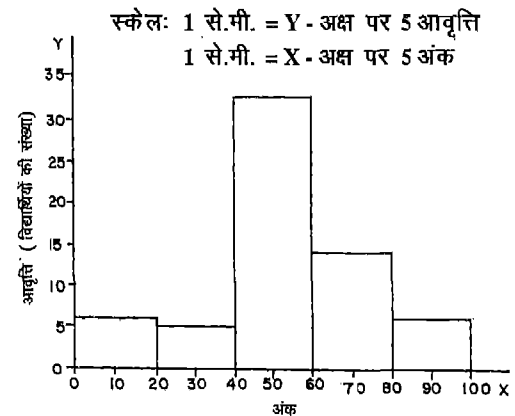
अंक (x) (वर्ग-अंतराल)	विद्यार्थियों की संख्या (आवृत्ति f)
0-20	6
20-40	5
40-60	33
60-80	14
80-100	6
कुल	**64**

आयत-चित्र का निर्माण करते समय हम निम्नलिखित बिंदुओं पर ध्यान देते हैं :

क. वर्ग-अंतरालों को x-अक्ष पर दिखाया जाना चाहिए।

ख. y -अक्ष पर आवृत्तियाँ दिखाई जानी चाहिए।

ग. दोनों अक्षों पर स्पष्ट लेबल लगाए जाने चाहिए तथा मापन के स्केल स्पष्ट रूप से बताए जाने चाहिए।

घ. सभी वर्ग-अंतरालों पर इस प्रकार आयत बनाए जाने चाहिए कि **प्रत्येक आयत का क्षेत्रफल उस वर्ग-अंतराल की आवृत्ति के अनुपात में हो।**

आरेख 5.4 : गणित में अंकों का आयात-चित्र

सारणी 5.1 के आवृत्ति वितरण को आरेख 5.4 के आयत-चित्र में दिखाया गया है। पहले आयत की ऊँचाई 6 है तथा वर्ग-अंतराल 0-20 है (उस वर्ग में आवृत्ति के बराबर) । वर्ग-अंतराल की चौड़ाई 20 है। अतः आयत का क्षेत्रफल 20 x 6=120 है। इसी प्रकार दूसरे आयत की ऊँचाई 5 तथा चौड़ाई 20 है। अतः दूसरे आयत का क्षेत्रफल 20 x 5=100 है। हम देखते हैं कि प्रत्येक आयत का क्षेत्रफल उस वर्ग-अंतराल में आवृत्ति से 20 गुना अधिक है। गुणनखंड 20 को **आनुपातिकता स्थिरांक** (constant of proportionality) कहते हैं।

संपूर्ण आयत-चित्र का क्षेत्रफल कुल आवृत्तियों के अनुपात में है (20 x 64)।

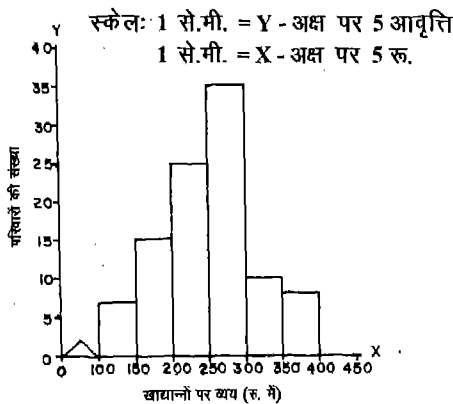
उदाहरण 2 : समान वर्ग अंतराल

किसी शहर में 100 परिवारों का खाद्य-सामग्रियों पर मासिक व्यय निम्नलिखित है :

सारणी - 5.2

100 परिवारों का खाद्य-सामग्रियों पर मासिक व्यय

खाद्यान्नों पर मासिक व्यय (रु. में) x	परिवारों की संख्या f
100-150	7
150-200	15
200-250	25
250-300	35
300-350	10
350-400	8
कुल	**100**

आरेख 5.5 में आयत-चित्र दिखाया गया है। हम देख सकते हैं कि x-अक्ष पर एक बल (kink) है। ऐसा इसलिए है, कि हम मूल-बिंदु 0 व्यय से बढ़कर सीधे 100 व्यय पर पहुँचते हैं। अतः 100 को भी मूल बिंदु चुना जा सकता है।

आरेख 5.5 : खाद्यानों पर मासिक व्यय का आयात-चित्र

उदाहरण 3 : असमान वर्ग अंतराल

निम्नलिखित सारणी परिवारों के प्रति-व्यक्ति मासिक व्यय का आवृत्ति वितरण दिखाती है :

सारणी - 5.3

प्रति व्यक्ति मासिक व्यय का आवृत्ति वितरण

प्रति व्यक्ति पारिवारिक मासिक व्यय (रु. में)	परिवारों की संख्या f	आयतों की ऊँचाई f' = f ÷ वर्ग अंतरालों की चौड़ाई
0 - 15	161	≈ 11
15 - 25	152	≈ 15
25 - 50	60	≈ 2
50 - 100	27	≈ 1
कुल	**400**	

चूँकि वर्ग अंतराल असमान हैं, अतः हम निम्न दो विधियों में से किसी एक का प्रयोग कर सकते हैं।

विधि 1.

पहला चरण : f' को प्राप्त करने के लिए प्रत्येक वर्ग में आवृत्ति को वर्ग-अंतराल की चौड़ाई से भाग दें। इससे हमें प्रत्येक इकाई वर्ग-अंतराल की आवृत्ति प्राप्त होती है।

दूसरा चरण : y-अक्ष पर f' को तथा x-अक्ष पर वर्ग अंतरालों को मापें। आयत की ऊँचाई $f' = \frac{f}{h}$ है, वर्ग अंतराल की चौड़ाई h है तथा आयत का क्षेत्रफल f है। **अतः आयत का क्षेत्रफल वर्ग-अंतराल में आवृत्तियों के बराबर होता है।**

विधि 2.

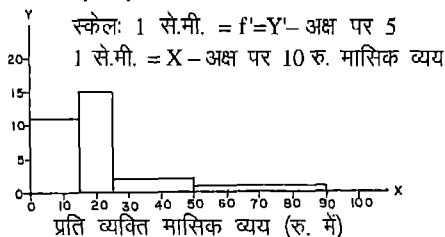
हम सबसे कम चौड़ाई वाले आयतों को चुनते हैं। सबसे छोटे वर्गों की आवृत्तियों को ज्यों-का-त्यों रहने दें। अन्य वर्गों में आवृत्तियों को अनुपात में निकालें, जैसा नीचे दिखाया गया है –

0-15	$\frac{10}{15}$ x 161 ≈ 107
15-25	$\frac{10}{10}$ x 152 ≈ 152
25-50	$\frac{10}{25}$ x 60 ≈ 24
50-100	$\frac{10}{50}$ x 27 ≈ 5

आवृत्तियों के साथ समायोजन कर एक आयत-चित्र बनाएँ। (आरेख 5.6)

(यह विधि संतोषजनक नहीं है, क्योंकि समायोजित आवृत्ति कुल प्रेक्षित आवृत्तियों के बराबर नहीं है)

उदाहरण 4 : समावेशित वर्ग अंतराल

124 कंपनियों के स्टॉक-विक्रय अनुपात नीचे सारणी में दिए गए हैं :

आरेख 5.6 : आयत चित्र (असमान वर्ग अंतराल)

सारणी - 5.4

स्टॉक विक्रय का आवृत्ति वितरण

स्टॉक विक्रय अनुपात (प्रतिशत)	कंपनियों की संख्या
0.1-5.0	2
5.1-10.0	3
10.1-15.0	8
15.1-20.0	14
20.1-25.0	38
25.1-30.0	59
कुल	124

इन आँकड़ों के आधार पर एक आयत-चित्र बनाएँ।

हम देखते हैं कि ये वर्ग-अंतराल **समावेशित विधि** द्वारा बनाए गए हैं (जहाँ उच्च तथा निम्न दोनों सीमाएँ वर्ग-अंतरालों में शामिल हैं) । आयत चित्र बनाने के लिए हमें दोनों वर्गों के बीच की दूरी भरने की आवश्यकता होगी। इसके लिए वर्गों को समायोजित करना होगा। समायोजित वर्ग अंतरालों के साथ आवृत्ति वितरण नीचे सारणी में दिए गए हैं —

सारणी - 5.5

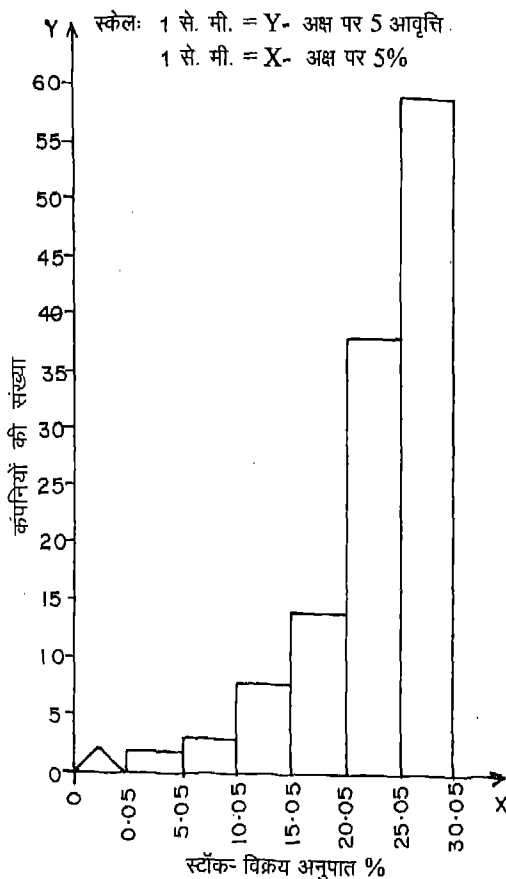
समायोजित वर्ग-अंतरालों सहित स्टॉक-विक्रय अनुपात का आवृत्ति वितरण

स्टॉक विक्रय अनुपात (%) समायोजित वर्ग-सीमाएँ	कंपनियों की संख्या
0.05–5.05	2
5.05–10.05	3
10.05–15.05	8
15.05–20.05	14
20.05–25.05	38
25.05–30.05	59
कुल	124

चूँकि पहले वर्ग की उच्चतम वर्ग-सीमा 3 5.0 है तथा दूसरे वर्ग की न्यूनतम सीमा 5.1 है, अतः समायोजित सीमा $\frac{5.0+5.1}{2}$ = 5.05 । वस्तुतः हमने इस अंतर (5.1-5.0) के आधे भाग को उच्चतम सीमा में जोड़ दिया है तथा इसे न्यूनतम सीमा से घटा दिया है।

आरेख 5.7 में आयत-चित्र दिखाया गया है।

उदाहरण 5 : आवृत्ति वितरण, जब केवल वर्ग-अंतरालों के मध्य बिंदु दिए गए हों।

मान लें कि x के आवत्ति वितरण निम्नलिखित हैं

आरेख 5.7 : स्टॉक-विक्रय अनुपात के वितरण का आयात-चित्र

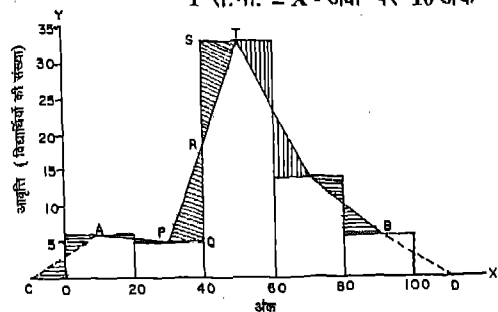
सारणी - 5.6

वर्ग - अंतरालों के मध्य मान x	आवृत्ति f
10	4
20	15
30	25
40	18
50	8
कुल	7

पहले उन वर्ग-अंतरालों की जानकारी कर लेनी आवश्यक है, जिनके मध्य मान ऊपर दिए गए हैं। ये वर्ग-अंतराल हैं 5-15, 15-25, 25-35, 35-45, 45-55 । अब पहले की भाँति आयत-चित्र बनाएँ।

4. आवृत्ति बहुभुज (Frequency Polygon)

आवृत्ति बहुभुज को संलग्न आयतों के मध्य-बिंदुओं को सरल रेखा द्वारा जोड़ते हुए दिखाया जाता है। इसे हम सारणी 5.1 में दिए गए आँकड़ों के लिए आवृत्ति बहुभुज का निर्माण कर दिखा सकते हैं। इसके लिए आरेख 5.4 में आयत-चित्र दिया गया है। सुविधा के लिए हमने आरेख 5.8 में एक बार फिर आयत चित्र बनाया है।

आवृत्ति बहुभुज बनाते समय, हम देखते हैं कि कुछ ऐसे क्षेत्र छोड़ दिए गए हैं, जो आयत चित्र के भीतर हैं। कुछ ऐसे क्षेत्र जो आयत चित्र के अंदर नहीं थे, आवृत्ति बहुभुज के भीतर आ गए हैं। उदाहरण के लिए आरेख 5.8 देखें। त्रिभुज RST का छायांकित क्षेत्र, जो आयत-चित्र के भीतर था, बहुभुज के अंदर नहीं है। यह क्षेत्र बहुभुज के बाहर है। परंतु त्रिभुज PQR का छायांकित क्षेत्र बहुभुज में शामिल है, जो आयत चित्र में नहीं था। इसी प्रकार अन्य क्षेत्रों को देखें। हर बार बहुभुज के

आरेख 5.8 : आवृत्ति बहुभुज

भीतर कोई ऐसा क्षेत्र आ जाता है जो आयत चित्र के अंदर नहीं है।

अतः यह आवश्यक है कि आयत चित्र के बाहर छूटा हुआ कुल क्षेत्र आवृत्ति-बहुभुज के भीतर के क्षेत्र के बराबर हो। यह वर्गों के सही चुनाव द्वारा हो सकता है।

ध्यान देना चाहिए कि पहले वर्ग-अंतराल पर बने आयत के शीर्ष का मध्य-बिंदु A आवृत्ति बहुभुज का आरंभिक बिंदु नहीं बन सकता। आयत चित्र में A के बायीं ओर के छूटे हुए क्षेत्र की क्षतिपूर्त्ति के रूप में बहुभुज के भीतर कोई क्षेत्र शामिल नहीं किया गया है। इसी प्रकार B भी अंतिम बिंदु नहीं हो सकता। इस समस्या का हल हम निम्नलिखित रूप से निकालते हैं :

क. पहले वर्ग अंतराल (0 - 20) की चौड़ाई को बायीं ओर भी उतना ही बढ़ाना होगा, जैसा चित्र में दिखाया गया है।

ख. बढ़ाए गए वर्ग-अंतराल का मध्य-बिंदु C लें।

ग. C तथा A को एक सरल रेखा द्वारा मिलाएँ।

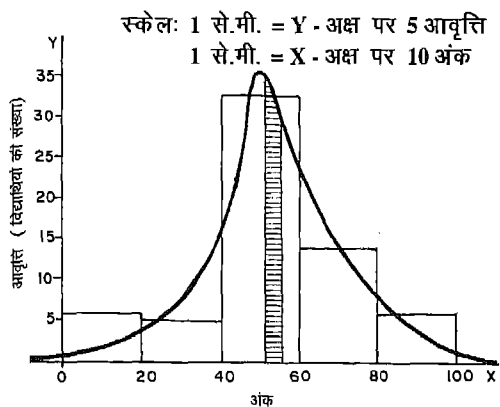
घ. अंतिम वर्ग अंतराल (30 - 100) की चौड़ाई दायीं ओर भी उतनी ही बढ़ाएँ, जैसा चित्र में दिखाया गया है।

ङ. बढ़ाए गए अंतराल का मध्य बिंदु D लें।

च. D को B के साथ एक सरल रेखा द्वारा मिलाएँ।

अतः, **आवृत्ति बहुभुज C से D तक है।**

5. आवृत्ति वक्र (Frequency Curve)

जैसा आरेख 5.9 में दिखाया गया है आवृत्ति वक्र आवृत्ति बहुभुज के बिंदुओं से होकर जाने वाला **मुक्तहस्त से बनाया गया एक ऐसा वक्र** है जो आवृत्ति बहुभुज के यथासंभव निकटतम बिंदुओं से होकर गुजरता है। (ये बिंदु आयत चित्र की आयतों के शीर्ष के मध्य भाग पर स्थित होते हैं।) यह आवश्यक नहीं कि आवृत्ति वक्र आवृत्ति बहुभुज के सभी बिंदुओं से होकर गुजरे। परंतु, यह उनके जितना संभव हो सके उतने निकट से होकर गुजरता है। आवश्यक है कि —

आरेख 5.9 : आवृत्ति वक्र (अंकों के बहुभुज पर अध्यारोपित)

सारणी - 5.6

हम सारणी-5.1 के गणित के अंकों के आवृत्ति वितरण को देखें

अंक x	विद्यार्थियों की संख्या f	'से कम' संचयी आवृत्ति	'से अधिक' संचयी आवृत्ति
0-20	6	6 = 6	64 = 6+5+33+14+6
20-40	5	6+5 = 11	58 = 5+33+14+6
40-60	33	6+5+33 = 44	53= 33+14+6
60-80	14	6+5+33+14 = 58	20 = 14+6
80-100	6	6+5+33+14+6 = 64	6 = 6
कुल	**64**		

क. **किसी भी वर्ग अंतराल में आवृत्ति वक्र के अन्दर का क्षेत्र, उस वर्ग अंतराल की आवृत्ति के अनुपात में होता है।**

ख. **आवृत्ति वक्र के अंदर का कुल क्षेत्र कुल आवृत्ति के अनुपात में होता है।**

आरेख-5.9 में सारणी-5.1 में दिए गए अंकों के वितरण को दिखाया गया है।

6. संचयी आवृत्ति वक्र — तोरण (Cumulative Frequency Curve — Ogive)

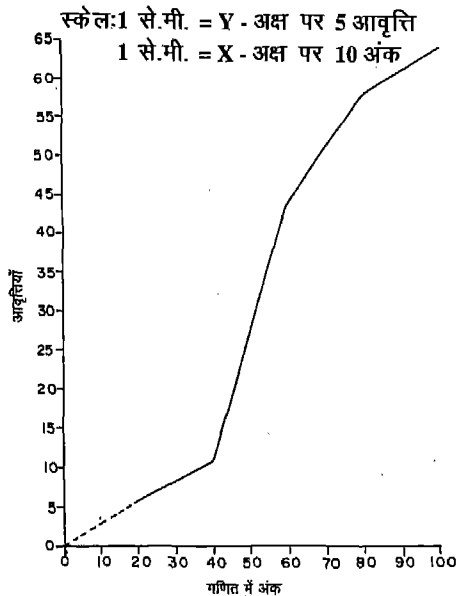
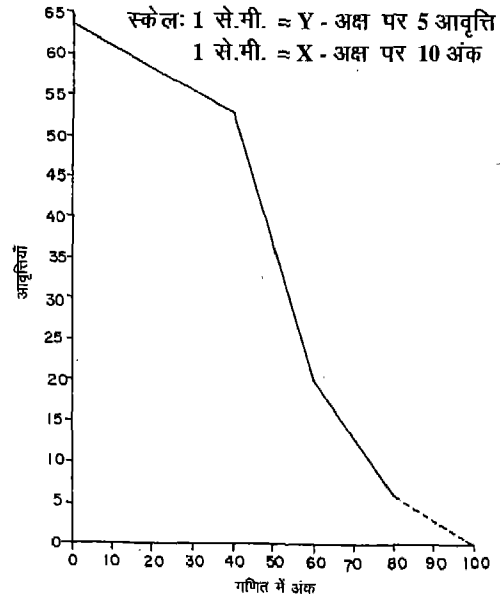
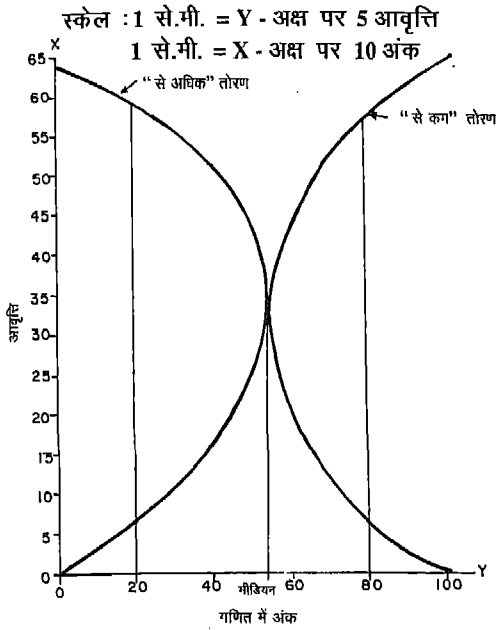
हम देखते हैं कि

6 विद्यार्थियों को 20 **'से कम'** अंक मिले।
11 विद्यार्थियों को 40 **'से कम'** अंक मिले।
44 विद्यार्थियों को 60 **'से कम'** अंक मिले।
58 विद्यार्थियों को 80 **'से कम'** अंक मिले।
64 विद्यार्थियों को 100 **'से कम'** अंक मिले।

अतः 6, 11, 44, 58 तथा 64 को **'से कम'** आवृत्ति कहते हैं तथा ये उपर्युक्त सारणी के तीसरे कॉलम में दिखाई गई हैं। हम इन 'से कम' संचयी आवृत्तियों को संगत वर्ग अंतरालों के उच्चतम अंतिम बिंदुओं तक अंकित करते हैं। अतः 6 को 20 के सामने 11 को 40 के सामने, 44 को 60 के सामने, 58 को 80 के तथा 64 को 100 के सामने अंकित करते हैं, जैसा आरेख 5.10 में दिखाया गया है।

इसी प्रकार हम 'से अधिक' संचयी आवृत्तियों को वर्ग अंतरालों की न्यूनतम सीमा तक अंकित करते हैं क्योंकि

64 विद्यार्थियों को 0 **'से अधिक'** अंक मिले हैं।
58 विद्यार्थियों को 20 **'से अधिक'** अंक मिले हैं।
53 विद्यार्थियों को 40 **'से अधिक'** अंक मिले हैं।
20 विद्यार्थियों को 60 **'से अधिक'** अंक मिले हैं।
6 विद्यार्थियों को 80 **'से अधिक'** अंक मिले हैं।
देखें आरेख- 5.11

यद्यपि आरेख 5.10 तथा 5.11 में हमने ग्राफ में बिंदुओं को सरल रेखा द्वारा मिलाया है, संचयी आवृत्ति-वक्र या तोरण इन बिंदुओं से होकर गुजरने वाले एक मुक्त हस्त वक्र द्वारा दिखाया जाता है।

आरेख 5.10 : 'से कम' संचयी आवृत्तियाँ

आरेख 5.11 : 'से अधिक' संचयी आवृत्तियाँ

आरेख 5.12 : 'से कम' तथा 'से अधिक' तोरण

आरेख 5.12 में इन दोनों वक्रों को एक साथ दिखाया गया है।

यहाँ हम देखते हैं कि —

'से कम' तोरण कभी घटता हुआ नहीं है तथा 'से अधिक' तोरण कभी बढ़ता हुआ नहीं है।

अतः x के किसी भी मान के लिए,

'से कम' तोरण की कोटियाँ ऐसे विद्यार्थियों की संख्या दर्शाती हैं जिन्हें x से कम अंक मिले हों, तथा **'से अधिक' तोरण की कोटियाँ ऐसे विद्यार्थियों की संख्या दर्शाती हैं जिन्हें x से अधिक अंक मिले हों।**

दोनों वक्रों के प्रतिच्छेद-बिंदु पर हमें x के मूल्य (अंकों की संख्या) इस प्रकार प्राप्त होते हैं ताकि x से अधिक या कम अंक प्राप्त करने वाले विद्यार्थियों की संख्या बराबर हो। 'x' के इस मूल्य को 'मध्यिका' (मीडियन) कहते हैं।

अभ्यास

1. दंड आरेख तथा आयत चित्र में अंतर बताएँ।
2. आयत चित्र बनाने की प्रक्रिया का वर्णन करें, जब वर्ग-अंतराल —
   क. समान हों
   ख. समान न हों।
3. निम्नलिखित के निर्माण की विधि का वर्णन करें
   क. आवृत्ति बहुभुज, तथा
   ख. आवृत्ति वक्र
4. क. आवृत्ति वक्र के अंतर्गत संपूर्ण क्षेत्र क्या दिखाता है?
   ख. आवृत्ति वक्र के अंतर्गत $x=\alpha$ पर कोटि की बायीं ओर का क्षेत्र क्या दिखाता है; जहाँ α संपूर्ण परास के अंतर का कुल क्षेत्र हो?
   ग. यदि $x=\alpha$ तथा $\beta=\alpha$, $\beta>\alpha$, ये परास के अंतर्गत x के दो मूल्य हैं, तो आवृत्ति वक्र के अंतर्गत कोटियों $x=\alpha$ तथा $x=\beta$ इन दोनों के बीच का क्षेत्र क्या दर्शाता है?
   घ. आवृत्ति-वक्र का एक आरेख बनाएँ तथा (क), (ख) तथा (ग) में दिए गए क्षेत्रों को दिखाएँ।

5. x का एक आवृत्ति वितरण नीचे दिया गया है:

वर्ग अंतराल	आवृत्ति
10 -20	15
20 -40	35
40 - 80	65
80 - 100	5
कुल	**120**

क. **'से कम'** संचयी आवृत्तियाँ निकालें।

ख. **'से अधिक'** संचयी आवृत्तियाँ निकालें।

ग. एक ग्राफ पर **'से कम'** तथा **'से अधिक'** तोरण बनाएँ।

घ. दोनों तोरण एक-दूसरे को कहाँ काटते हैं?

ङ. x का मध्यिका मूल्य निकालें।

6. 450 श्रमिकों की दैनिक मजदूरी आय (रु. में) का आवृत्ति वितरण नीचे दिया गया है :

दैनिक आय (रु. में)	व्यक्तियों की संख्या
70-80	44
80-90	120
90-100	80
100-110	76
110-120	50
120-130	45
130-140	25
140-150	10
कुल	**450**

क. एक आवृत्ति आयत-चित्र बनाकर उस पर एक आयत बहुभुज तथा एक आवृत्ति वक्र को अध्यारोपित करें।

ख. संचयी आवृत्तियाँ निकालें तथा दो तोरण बनाएँ।

ग. 'मध्यिका' दैनिक मजदूरी निकालें।

7. 80 सार्वजनिक क्षेत्र इकाइयों के क्षमता-प्रयोग का आवृत्ति वितरण नीचे दिया गया है —

क्षमता प्रयोग (%)	इकाइयों की संख्या
50 से कम	14
50.1-60	14
60.1-70	5
70.1-80	10
80.1-90	10
90.1-100	7
100 से अधिक	20
कुल	**80**

क. एक आयत चित्र बनाएँ तथा उस पर एक आवृत्ति बहुभुज तथा एक आवृत्ति वक्र को अध्यारोपित करें (ध्यान रहे कि उत्तरोत्तर वर्ग अंतरालों के अंतिम बिंदुओं के बीच अंतर है। हम वर्ग अंतरालों का अंतर दूर करने के लिए इस प्रकार सुधार सकते हैं : 50.05, 50.05-60.05, 60.05-70.05, ...100.50)।

ख. पहला तथा अंतिम वर्ग अंतराल यदि विवृत (open) हो तो क्या होगा? आप कैसे इस स्थिति का हल निकालेंगे?

अध्याय - 6

केंद्रीय प्रवृत्ति की माप

हम पहले के अध्यायों में आँकड़ों के वर्गीकरण, सारणीकरण तथा उनके आरेखी प्रस्तुतीकरण की चर्चा कर चुके हैं। ये आँकड़ों की तुलना करने तथा उन्हें समझने में मदद करते हैं तथा जनसंख्या की विशेषताओं के संबंध में भी निष्कर्ष निकालने में सहायक होते हैं।

इसके बाद, उन चरों के मूल्यों के वितरण के बारे में जानना उपयोगी होगा, जिनके बारे में आँकड़े एकत्र किए गए हैं। उदाहरण के लिए :

क. चर के सबसे अधिक तथा सबसे कम मूल्य क्या हैं तथा मूल्यों के वितरण का परास कितना है

ख. वितरण का वह केंद्रीय मूल्य कौन-सा है जिसके आसपास अन्य मूल्य वितरित हैं,

ग. केंद्रीय मूल्य के आसपास मूल्यों के वितरण का परिमाण कितना है,

घ. क्या चर के मूल्य परास के अंतर्गत समान रूप से वितरित हैं या परास की ऊपरी या निचली सीमा के निकट या उसके बीच में एकत्रित हैं?

इस अध्याय में, हम केंद्रीय प्रवृत्ति के मापों की चर्चा करेंगे तथा परिक्षेपण के मापों का वर्णन अगले अध्यायों में करेंगे।

केंद्रीय प्रवृत्ति की माप को **अवस्थिति की माप** (measures of location) या **औसत** भी कहते हैं।

1. औसत (Averages)

'औसत मूल्य' शब्द का प्रयोग दैनिक जीवन में होता है। यह वह एकल मान है जो दिए गए समुच्चय के सभी मूल्यों का प्रतिनिधित्व करता है। उदाहरण के लिए, आपकी कक्षा के विद्यार्थियों की औसत लंबाई कक्षा के सभी विद्यार्थियों की लंबाई का प्रतिनिधित्व करती है। किसी क्षेत्र के लोगों की औसत आय उस क्षेत्र के सभी लोगों की आय का प्रतिनिधित्व करती है। अलग-अलग स्थितियों में निम्नलिखित औसत उपयुक्त होते हैं :

क. समांतर माध्य

ख. मध्यिका

ग. बहुलक

घ. ज्यामितीय माध्य

ङ. हरात्मक माध्य

औसतों की विशेषताओं का वर्णन करने से पहले, हमें एक अच्छे औसत की आवश्यकताएँ जान लेनी चाहिए (यहाँ हम ज्यामितीय तथा हरात्मक माध्य की चर्चा नहीं करेंगे क्योंकि ये ग्यारहवीं कक्षा के पाठ्यक्रम का अंग नहीं हैं)।

2. एक अच्छे औसत की आवश्यकताएँ

एक अच्छे औसत के लिए आवश्यक है कि यह

क. **सुपरिभाषित हो।**

ख. **सभी मूल्यों पर आधारित हो।**

ग. **सरल एवं समझने में आसान हो।**

घ. **परिकलन में आसान हो।**

ङ. **बीजगणितीय विवेचन के योग्य हो।**

यदि औसत सुपरिभाषित न हो तो विभिन्न व्यक्ति इसके विभिन्न अर्थ निकाल सकते हैं। परिकलनों के द्वारा अलग-अलग परिणामों के प्राप्त होने की भी आशंका रहती है। फिर, यदि इसमें समुच्चय के सभी मूल्यों का समावेश न हो तो यह सभी मूल्यों का प्रतिनिधि नहीं हो सकता है। औसतों का बीजगणितीय विवेचन आगे के अध्ययन में उपयोगी रहेगा।

3. समांतर माध्य (Arithmetic Mean)

मान लें की आपकी कक्षा के दस विद्यार्थियों ने दसवीं कक्षा की परीक्षा में निम्नलिखित प्रतिशत अंक प्राप्त किए हैं :

65	59	75	79	85
45	55	70	77	72

अंकों को जोड़कर तथा उन्हें विद्यार्थियों की कुल संख्या से भाग देकर, समांतर माध्य निकाला जाता है, जैसे —

$$\text{समांतर माध्य} = \frac{682}{10} = 68.2$$

अंकों का औसत प्रतिशत 68.2 है।

समांतर माध्य की एक रोचक विशेषता यह है कि **समांतर माध्य से विचलनों के मूल्यों का योग शून्य होता है**। इसीलिए, समांतर माध्य वितरण का केंद्रीय मूल्य है।

उपर्युक्त उदाहरण के लिए, समांतर माध्य से अंकों का विचलन निम्नलिखित सारणी में दिखाया गया है :

सारणी - 6.1

समांतर माध्य से अंकों का विचलन

अंकों का प्रतिशत	**समांतर माध्य (=68.2) से विचलन**	
65	-3.2	
59	-9.2	
75		+6.8
79		+10.8
85		+16.8
45	-23.2	
55	-13.2	
70		+1.8
77		+8.8
72		+3.8
कुल	**-48.8**	**+48.8**

हमने देखा कि समांतर माध्य से धनात्मक विचलनों का योग इससे ऋणात्मक विचलनों के योग के बराबर है। दूसरे शब्दों में धनात्मक तथा ऋणात्मक विचलन एक-दूसरे को निष्क्रिय कर देते हैं। इस दृष्टि से, समांतर माध्य भौतिकी के 'गुरुत्व केंद्र' की आवधारणा के समान है।

उदाहरण 1

क. इसकी जाँच की जा सकती है कि किसी अन्य मूल्य से विचलनों का योग शून्य नहीं होता।

ख. 'शून्य' से विचलनों का योग कितना होगा? (शून्य से विचलन मूल बिंदु से विचलन भी कहे जाते हैं)।

सामान्यतः, मान लें कि $x_1, x_2, ...x_n$ किसी चर x के मूल्य हैं, तो उनका समांतर माध्य है

$$\frac{x_1 + x_2 + ... + x_n}{n}$$

जिसे सांकेतिक रूप से $\bar{x}$ द्वारा दिखाते हैं। इसे हम संक्षेप में इस प्रकार वर्णित करते हैं :

$$\bar{x}=\frac{1}{n}\sum x_i$$

जहाँ $\sum x_i$ का अर्थ n मूल्यों का योग होता है। '$\sum$' एक ग्रीक 'वर्ण' है जिसे 'सिग्मा' कहते हैं। यह सिग्मा का बड़ा अक्षर है।

उदाहरण 2

जाँच करें कि

$$\sum(x_i-\bar{x}) = 0$$

जिसका अर्थ है कि समांतर माध्य से विचलनों का योग शून्य है।

इसे हम इस प्रकार लिखें :

$$\sum(x_i-\bar{x})=(x_1-\bar{x})+(x_2-\bar{x})+......+(x_n-\bar{x})$$

$$=(x_1+x_{2+}......{}_+x_n)-n\bar{x}$$

$$= 0$$

क्योंकि $\bar{x}=\frac{1}{n}\sum x_i$

उदाहरण 3

5 परिवारों के मासिक आय (रु. में) निम्नलिखित है :

1550	1715	1690	820	1150

समांतर माध्य 1385 है, जिसका अर्थ है कि इन परिवारों का औसत मासिक व्यय 1385 रु. है।

ऐसा भी हो सकता है कि समांतर माध्य समुच्चय के किसी भी मूल्य के बराबर न हो।

4. समांतर माध्य की विशेषताएँ

अब हम समांतर माध्य की विशेषताओं को देखेंगे ताकि इस माध्य की तुलना अन्य केंद्रीय प्रवृत्तियों के साथ करने में सुविधा हो।

क. जैसा पहले बताया गया है कि समांतर माध्य से मूल्यों के विचलनों का योग शून्य है।

ख. मान लें कि हम कुछ परिवारों की मासिक आय लेते हैं। समांतर माध्य के द्वारा यह मालूम होगा कि यदि आय को सभी परिवारों के बीच समान रूप से वितरित किया जाए तो प्रत्येक परिवार की आय क्या होगी?

ग. समांतर माध्य चर का एक सरल फलन (फंक्शन) है। इसके दो लाभ हैं :

i. समांतर माध्य की अंकगणितीय विशेषताओं को निकालना आसान है।

ii. दिए गए आँकड़ों के आधार पर स्पष्ट रूप से इसका परिकलन किया जा सकता है।

घ. समांतर माध्य के परिकलन में चर के सभी मूल्यों का समान महत्त्व है। ऐसा इसलिए है कि चर का बहुत बड़ा या बहुत छोटा मूल्य भी समांतर माध्य को प्रभावित करता है।

उदाहरण के लिए, किसी इलाके के 5 परिवारों के दैनिक व्यय (रु.में) निम्नलिखित हैं :

25	28	32	27	33

इन मूल्यों का समांतर माध्य 29 है।

अब मान लें कि कोई धनी व्यक्ति उस इलाके में आ जाता है जिसका दैनिक व्यय 125 रु. है। यदि हम समांतर माध्य को फिर से निकालें तो इसका मूल्य 45 रु. होगा। चूँकि समांतर माध्य में 50 प्रतिशत से अधिक की वृद्धि हो गई है, अतः ऐसा सोचा जा सकता है कि पूरा इलाका ही धनी हो गया है। हालाँकि, हम यह भी देखते हैं कि 6 में से 5 परिवार पहले की ही तरह व्यय कर रहे हैं। अतः समांतर माध्य के मूल्य में वृद्धि नए आए हुए व्यक्ति के कारण ही हुई है।

5. मध्यिका (Median)

समांतर माध्य वितरण का केंद्रीय माप है क्योंकि समांतर माध्य के धनात्मक तथा ऋणात्मक विचलन एक-दूसरे को बराबर कर देते हैं। दूसरी ओर, मध्यिका वितरण का केंद्रीय माप इसलिए है कि

मध्यिका से नीचे तथा ऊपर वाले मूल्यों की संख्या एक-दूसरे के बराबर होती है। इसका अर्थ है कि मध्यिका से नीचे के मूल्यों की संख्या मध्यिका से ऊपर के मूल्यों की संख्या के बराबर होती है।

मान लें कि आपकी कक्षा में 17 विद्यार्थी हैं। उन्होंने कक्षा दसवीं की परीक्षा में पिछले वर्ष निम्नलिखित प्रतिशत अंक प्राप्त किए :

35	63	61	29	54	46	48	57	43
60	39	40	51	68	38	56	53	

यदि हम अंकों को आरोही क्रम में रखें तो हमें निम्नलिखित संख्याएँ इस प्रकार प्राप्त होंगी :

29	35	38	39	40	43	46	48	51
53	54	56	57	60	61	63	68	

इसका अर्थ है कि 8 विद्यार्थियों ने 51 प्रतिशत से कम तथा 8 विद्यार्थियों ने इससे अधिक अंक प्राप्त किए। अतः मध्यिका = 51 प्रतिशत। मध्यिका तथा समांतर माध्य में अंतर है। समांतर माध्य में विचलनों का संख्यात्मक मूल्य संतुलन को बनाए रखता है। परंतु मध्यिका में, मध्यिका से ऊपर की संख्याएँ, इससे नीचे की संख्याओं को संतुलित करती हैं।

सामान्यतः यदि x के n मूल्य हैं, तो उन्हें आरोही क्रम में इस प्रकार लगाया जा सकता है :

$x_1 < x_2 < \ldots\ldots, < x_n$

मान लें कि यदि *n* **विषम** है, तो मध्यिका $= \left(\frac{n+1}{2}\right)$ वाँ मूल्य।

हालाँकि, यदि *n* **सम** हो, तो दो मध्य-मूल्य हो जाते हैं, जैसे —

$\left(\frac{n}{2}\right)$ वाँ मूल्य तथा $\left(\frac{n}{2}\right)+1$ वाँ मूल्य.......

ऐसी स्थिति में, मध्यिका

$$= \frac{\left(\frac{n}{2}\right)\text{वाँ मूल्य} \left(\frac{n}{2}+1\right)\text{ वाँ मूल्य}}{2}$$

जो दोनों मध्य-मूल्यों का समांतर मांध्य है।

उदाहरण 4

आपकी कक्षा के 6 विद्यार्थियों की लंबाई (से.मी. में) 160, 157, 156, 161, 159 तथा 162 है। मध्यिका लंबाई कितनी होगी?

यदि हम लंबाई को आरोही क्रम में रखें, तो

156	157	159	160	161	162

तीसरा तथा चौथा मूल्य, दोनों मध्य-मूल्य हैं। मध्यिका $= \frac{1}{2}(159 + 160) = 159.5$ होगी। इसकी तुलना समांतर माध्य (=159.2 लगभग) से करें।

6. मध्यिका की विशेषताएँ

मध्यिका की प्रमुख विशेषताएँ निम्नलिखित हैं :

क. यह सरल और समझने में आसान है।

ख. समांतर माध्य की भाँति मध्यिका सुपरिभाषित है।

ग. यह अधिकतम या न्यूनतम मूल्यों द्वारा प्रभावित नहीं होती है।

घ. मध्यिका को किसी दिए गए समुच्चय के मूल्यों के एक सरल फलन के रूप में नहीं दिखाया जा सकता है।

7. बहुलक (Mode)

बहुलक चर का ऐसा मान है जो सबसे अधिक बार घटित होता है।

आपके स्कूल की क्रिकेट टीम में विद्यार्थियों की उम्र की आवृत्ति सरणी नीचे दी गई है :

सारणी - 6.2

विद्यार्थियों की आयु की आवृत्ति सरणी

आयु	विद्यार्थियों की संख्या
14	2
15	2
16	4
17	2
18	1
कुल	**11**

चार विद्यार्थी ऐसे हैं जिनकी उम्र 16 वर्ष है। चूँकि इस आयु की आवृत्ति सबसे अधिक बार होती है, अतः बहुलक मूल्य 16 वर्ष है।

कभी-कभी बहुलक समांतर माध्य तथा मध्यिका से अधिक सार्थक औसत होता है। निम्नलिखित उदाहरण लें।

उदाहरण 5

कोई कंपनी जो वयस्कों के लिए जूते बनाती है, सिर्फ एक ही साइज के जूते बनाने में रुचि रखती है। यह आकार कौन-सा हो? इसके लिए कंपनी ने 2000 ग्राहकों का सर्वेक्षण उनके जूतों के आकार जानने के लिए किया जिससे निम्नलिखित आवृत्ति वितरण प्राप्त हुआ :

सारणी - 6.3

जूतों के आकार का आवृत्ति वितरण

जूतों के आकार	ग्राहकों की संख्या
5	55
6	164
7	360
8	528
9	410
10	260
11	137
12	86
कुल	**2000**

इस उदाहरण में न तो समांतर माध्य और न ही मध्यिका उपयुक्त औसत होगा। यहाँ बहुलक का प्रयोग उपयोगी होगा। चूँकि सबसे बड़ी आवृत्ति 528 है, इसलिए बहुलक 8 है। अतः कंपनी के लिए 8 नम्बर के जूते बनाना उपयोगी होगा।

8. बहुलक की विशेषताएँ

क. समांतर माध्य तथा मध्यिका की भाँति, बहुलक को आसानी से समझा जा सकता है।

ख. बहुलक अधिकतम और निम्नतम मानों से प्रभावित नहीं होता, क्योंकि यह केवल सबसे अधिक आवृत्ति पर निर्भर करता है।

ग. मध्यिका की भाँति, बहुलक दिए गए मानों के एक सरल फलन की तरह नहीं दर्शाया जा सकता।

9. अवर्गीकृत आँकड़ों के लिए समांतर माध्य का परिकलन

प्रत्यक्ष विधि (Direct Method)

मान लें कि $x_1, x_2, \ldots\ldots x_n$ चर x के n मान हैं। ये मान वर्ग अंतरालों में वर्गीकृत नहीं किए गए हैं। अतः प्रत्यक्ष विधि द्वारा x का समांतर माध्य इस प्रकार निकाला जाता है

$$\bar{x}=\frac{1}{n}\sum x_i$$

जहाँ x के सारे मानों को जोड़कर तथा उनके योग को उनकी कुल संख्या से भाग दिया जाता है।

कल्पित माध्य विधि (Assumed Mean Method)

मान लें कि A कल्पित माध्य है। अतः हम A से x का विचलन निकालते हैं, जैसे — $d_1 = x_1 - A$, $d_2 = x_2 - A, \ldots\ldots d_n = x_n - A$ । इसके बाद, विचलनों का समांतर माध्य निकालें,

$$\bar{d}=\frac{\sum d_i}{n} \text{ तथा } \bar{x}=A+\bar{d}$$

दूसरें शब्दों में विचलनों (d) के समांतर माध्य निकालें तथा उसमें A को जोड़ें, जिसे पहले घटाया गया था।

पद-विचलन विधि (Step-deviation Method)

पहले कल्पित माध्य से विचलनों (d) को निकालें। विचलनों को एक समापवर्तक h से भाग दें :

$$d'_1=\frac{d_1}{h}, d'_2=\frac{d_2}{h}, ... d'_n=\frac{d_n}{h}$$

इससे हमें

$\bar{x}=A+h\bar{d}'$ प्राप्त होगा

जहाँ

$$\bar{d}'=\frac{1}{n}\sum d'_i \text{।}$$

दूसरे शब्दों में, समांतर माध्य $d'_1, d'_2, ... d'_n$ निकालें तथा इन्हें h से गुणा करें (क्योंकि इन्हें हमने पहले h से भाग दिया था) तथा इसमें A को जोड़ें (क्योंकि पहले हमने A को घटाया था) ।

उदाहरण 6

मान लें कि 5 व्यक्तियों की मासिक आय x (रु. में) नीचे दी गई है: 6550, 7550, 9550, 4550, 8000। इनका समांतर माध्य निकालें।

प्रत्यक्ष विधि

प्रत्यक्ष विधि में, हम सभी संख्याओं को जोड़ देते हैं तथा उन्हें व्यक्तियों की संख्याओं से भाग देते हैं। इससे हमें

$$\bar{x}=\frac{36200}{5}=7240$$

अतः 5 व्यक्तियों की आय का समांतर माध्य 7240 रु. है।

कल्पित माध्य विधि द्वारा समांतर माध्य का परिकलन

मान लें कि कल्पित माध्य A=5550 है। कल्पित माध्य से विचलन (d) निकालें, जैसे — d_1=1000, d_2=2000, d_3=4000, d_4–1000, तथा d_5=2450

ताकि, $\sum di = 8450$ तथा $\bar{d}= 1690$ तो $\bar{x} = A+\bar{d}$ के अनुसार $\bar{x} = 5550+1690 = 7240$ जो पहले की भाँति है।

पद-विचलन विधि

हम विचलनों (d) को एक समापवर्तक 1000 से भाग देकर $d'_1= 1$, $d'_2= 2$, $d'_3= 4$, $d'_4= -1$, तथा $d'_5= 2.45$ निकालते हैं।

$$\frac{\sum d'_i}{5}=\frac{8.45}{5}=1.69;$$

तथा पहले की तरह,

$$\bar{x} = A+1000\bar{d}'_1$$
$$= 5550 + 1000 \times 1.69$$
$$= 7240$$

यहाँ ध्यान देना चाहिए कि कल्पित मान से विचलन तथा पद-विचलन विधि दोनों ही परिकलन के बोझ को बहुत कम कर देते हैं। ये विशेषकर तब उपयोगी होते हैं जब संख्याएँ बहुत बड़ी होती हैं।

10. समांतर माध्य का परिकलन – आवृत्ति सरणी (array) के लिए

चर x एक असंतत चर है जिसके मान $x_1, x_2, ..., x_n$ हैं तथा आवृत्तियाँ $f_1, f_2, ..., f_n$ हैं (वर्ग-अंतराल नहीं दिए गए हैं) ।

आवृत्ति सरणी को हम इस प्रकार दिखा सकते हैं

सारणी - 6.4

आवृत्ति सरणी

x के मान (x)	आवृत्ति (f)
x_1	f_1
x_2	f_2
x_n	f_n
कुल	**N**

x के मान x_1 की आवृत्ति f_1 बार होती है, x के मान x_2 की आवृत्ति f_2 बार होती है, x के मान x_n की आवृत्ति f_n बार होती है।

अतः प्रत्यक्ष विधि द्वारा x का समांतर माध्य निम्नलिखित है :

$$\bar{d} = \frac{\overbrace{x_1+......+x_1}+\overbrace{x_2+......+x_2}+\overbrace{......x_n+......x_n}}{N}$$

$$= \frac{f_1x_1+f_2x_2+........+f_nx_n}{N} = \frac{\sum f_ix_i}{N} = \frac{\sum f_ix_i}{\sum f}$$

कल्पित माध्य विधि द्वारा समांतर माध्य का परिकलन

मान लें कि A कल्पित माध्य है। कल्पित माध्य से x के मानों का विचलन निकालें

$d_1 = x_1 - A, ..., d_n = x_n - A$

इसके बाद,

$$\bar{d} = \frac{1}{N}\sum f_i d_i = \frac{1}{\sum f}\sum f_i d_i$$

$$\bar{x} = A + \bar{d}$$

पद-विचलन विधि द्वारा समांतर माध्य का परिकलन

मान लें कि h समापवर्तक है। प्रत्येक विचलन (d) को h द्वारा भाग देकर

$d'_1 = \frac{d_1}{h}, ..., d'_n = \frac{d_n}{h}$ निकालें।

उसके बाद

$$\bar{d}' = \frac{1}{N}\sum f_i d_i = \frac{1}{\sum f}\sum f_i d_i$$

तथा

$$\bar{x} = A + h\bar{d}'$$

दूसरे शब्दों में, हम $\bar{d}'$ को h से गुणा करते हैं (क्योंकि पहले इसे h से भाग दिया गया था) ।

उदाहरण 7

100 परिवारों का एक सर्वेक्षण उनके परिवारों का आकार जानने के लिए किया गया। सर्वेक्षण के परिणाम नीचे सारणी में आवृत्ति-सरणी के रूप में वर्गीकृत किए में गए है :

सारणी - 6.5

परिवारों के आकार की आवृत्ति सरणी

परिवारों के आकार x	परिवारों की संख्या f
1	5
2	15
3	25
4	35
5	10
6	5
7	3
8	2
कुल	**$\sum f = 100$**

परिवारों के आकार का समांतर माध्य निकालें।

प्रत्यक्ष विधि

पहले परिकलनों को निम्नलिखित सारणी में सारणीबद्ध करें।

सारणी - 6.6

समांतर माध्य का परिकलन
(प्रत्यक्ष विधि द्वारा)

x	f	fx
1	5	5
2	15	30
3	25	75
4	35	140
5	10	50
6	5	30
7	3	21
8	2	16
कुल	**$\sum f = 100$**	**367**

अतः समांतर माध्य है

$$\bar{x} = \frac{367}{100} = 3.67$$

कल्पित माध्य विधि (Assumed Mean Method)

मान लें कि X का कल्पित माध्य 4 है। परिकलनों को नीचे सारणी में दिया गया है।

सारणी - 6.7

समांतर माध्य का परिकलन
(कल्पित माध्य विधि द्वारा)

x	d = x – 4	f	fd
1	-3	5	-15
2	-2	15	-30
3	-1	25	-25
4	0	35	0
5	1	10	10
6	2	5	10
7	3	3	9
8	4	2	9
	$\sum f = 100$		–70+37= –33

अतः $\bar{d} = -\frac{33}{100} = -0.33$ तथा $\bar{x} = 4 + \bar{d} = 4 - .33$ = 3.67 जो पहले की तरह है।

इस उदाहरण में पद-विचलन विधि के प्रयोग से कोई विशेष लाभ नहीं है क्योंकि X के मान छोटे ही हैं तथा कल्पित मान से विचलन सुविधाजनक है। एक दूसरा उदाहरण लें।

उदाहरण 8

व्यक्तियों की मासिक आय की आवृत्ति सारणी निम्नलिखित है।

सारणी - 6.8

आय की आवृत्ति सरणी

आय (रु. में)	व्यक्तियों की संख्या
11,500	629
12,000	1705
12,500	1800
13,000	725
13,500	420
14,000	127
14,500	85
15,000	9
कुल	**5500**

प्रत्यक्ष विधि

परिकलनों को निम्नलिखित सारणी में रखें।

सारणी - 6.9

समांतर माध्य का परिकलन
(प्रत्यक्ष विधि द्वारा)

आय	व्यक्तियों की संख्या	fx
11,500	629	7,233,500
12,000	1705	20,460,000
12,500	1800	22,500,000
13,000	725	9,425,000
13,500	420	5,670,000
14,000	127	1,778,000
14,500	85	1,232,500
15,000	9	135,000
कुल	**5,500**	**68,434,000**

अतः आय का समांतर माध्य है

$$\frac{\sum f_i x_i}{N} = \frac{68,434,000}{5,500} \approx 12,442.55$$

कल्पित माध्य विधि

मान लें कि कल्पित माध्य 12,500 है। परिकलनों को निम्नलिखित सारणी में व्यवस्थित करें।

सारणी - 6.10

समांतर माध्य का परिकलन
(कल्पित माध्य विधि द्वारा)
(कल्पित माध्य =12,500)

आय	विचलन d=x – 12500	f	fd
11,500	-1000	629	-629,000
12,000	-500	1705	-852,500
12,500	0	1800	0
13,000	500	725	362,500
13,500	1000	420	420,000
14,000	1500	127	190,500
14,500	2000	85	170,000
15,000	2500	9	22,500
कुल		**5,500**	**-316,000**

इसलिए

$$\bar{d} = \frac{1}{N}\sum f_i d_i = \frac{316000}{5500} \approx -5745$$

तथा

$$\bar{x} = A + \bar{d} = 12442.55$$

पद-विचलन विधि

अब हम आय का समांतर माध्य पद-विचलन विधि द्वारा निकालें। परिकलनों को निम्नलिखित सारणी में दिखाया गया है :

सारणी - 6.11

समांतर माध्य का परिकलन
(पद-विचलन विधि द्वारा)
(कल्पित माध्य =12500 तथा समापवर्तक 500)

आय x	विचलन d=x–12500	$d' = \frac{d}{500}$	f	fd'
11500	-1000	-2	629	-1258
12000	-500	-1	1705	-1705
12500	0	0	1800	0
13000	500	1	725	725
13500	1000	2	420	840
14000	1500	3	127	381
14500	2000	4	85	340
15000	2500	5	9	45
कुल			**$\sum f = 5500$**	**-632**

अतः

$$\bar{d'} = \frac{1}{N}\sum f_i d_i' = -\frac{632}{5500} \approx -0.1149$$

तथा

$$\bar{x} = A + h\bar{d'}$$
$$= 12500 - 500 \times 0.1149$$
$$= 12442.55$$

यहाँ हमने देखा कि पद-विचलन विधि बड़े परिकलनों को काफी सरल बना देती है।

11. समांतर माध्य का परिकलन — वर्गीकृत आँकड़ों के लिए

हम मान लेते हैं कि x एक संतत चर है तथा x के मान **समान** वर्ग-अंतरालों में वर्गीकृत किए गए हैं। इन परिकलनों को निम्नलिखित उदाहरण द्वारा स्पष्ट किया जा सकता है।

उदाहरण 9

100 ग्रामीण परिवारों के मासिक व्यय का वितरण नीचे दिया गया है :

सारणी - 6.12

100 परिवारों के मासिक व्यय का आवृत्ति वितरण

मासिक व्यय रु. में (वर्ग अंतराल)	परिवारों की संख्या
75-125	10
125-175	22
175-225	38
225-275	16
275-325	8
325-375	4
375-425	2
कुल	**100**

मासिक व्यय के समांतर माध्य का परिकलन निम्नलिखित विधियों द्वारा करें :

क. प्रत्यक्ष विधि,

ख. कल्पित माध्य विधि, तथा

ग. पद-विचलन विधि ।

क. प्रत्यक्ष विधि

परिकलनों को निम्नलिखित सारणी में सारणीबद्ध करें :

सारणी - 6.13

समांतर माध्य का परिकलन
(प्रत्यक्ष विधि द्वारा)

वर्ग-अंतराल	वर्ग-अंतरालों के माध्य मान x	f	fx
1	2	3	4
75-125	100	10	1000
125-175	150	22	3300
175-225	200	38	7600
225-275	250	16	4000
275-325	300	8	2400
325-375	350	4	1400
375-425	400	2	800
कुल		**$\sum f = 100$**	**20500**

माध्य मान निम्नलिखित विधि द्वारा निकालें जाते है :

$\frac{75+125}{2}=100, \frac{125+175}{2}=150, \frac{175+225}{2}=200$

इत्यादि।

अतः समांतर माध्य $=\frac{20500}{100}=205$

ख. कल्पित माध्य विधि

मान लें कि कल्पित माध्य 200 है। उसके बाद समांतर माध्य निम्नलिखित सारणी में दिखाए गए है :

सारणी - 6.14

समांतर माध्य का परिकलन
(कल्पित माध्य विधि)
(कल्पित माध्य =200)

वर्ग अंतराल	मध्य मान x	विचलन 2 = x – 200	f	fd
1	2	3	4	5
75-125	100	-100	10	-1000
125-175	150	-50	22	-1000
175-225	200	0	38	0
225-275	250	50	16	800
275-325	300	100	8	800
325-375	350	150	4	600
375-425	400	200	2	400
कुल			**∑f = 100**	**500**

अतः

$$\bar{d}=\frac{\sum f_i d_i}{N}=\frac{\sum f_i d_i}{\sum f}=\frac{500}{100}=5$$

तथा $\bar{x}=A+\bar{d}=200+5=205$

ग. पद-विचलन विधि

ऊपर दी गई सारणी के कॉलम (2) को देखकर हम यह जान सकते हैं कि समापवर्तक 50 है। वस्तुतः यही वर्ग-अंतराल भी है। परिकलन नीचे सारणी में दिए गए हैं।

सारणी - 6.15

समांतर माध्य का परिकलन
(पद-विचलन विधि)
(कल्पित माध्य=200 तथा समापवतर्वक=50)

वर्ग-अंतराल	मध्य मान	d=x–200	$d'=\frac{d}{50}$	f	fd'
1	2	3	4	5	6
75-125	100	-100	-2	10	-20
125-175	150	-50	-1	22	-22
175-225	200	0	0	38	0
225-275	250	50	1	16	16
275-325	300	100	2	8	16
325-375	350	150	3	4	12
375-425	400	200	4	2	8
कुल				**∑f = 100**	**10**

अतः

$$\bar{d'}=\frac{\sum f_i d_i'}{N}=\frac{10}{100}=\frac{1}{10}=0.1$$

तथा

$$\bar{x}=A+h\bar{d}'=200+50\times0.1=205$$

जहाँ

A=200 तथा समापवर्तक h=50

12. मध्यिका का परिकलन

चरों के मूल्यों के समुच्चय को आरोही (या अवरोही) क्रम में लगाते हैं। मध्यिका मध्य मान को कहते हैं। यदि मानों की संख्या विषम है तो मध्यिका $\frac{101+1}{2}$=51 वाँ मान होगी। सामान्य रूप से यदि N = 2m+1 है, तो मध्यिका (m+1) वाँ मान होगा। अन्यथा, यदि n=2m सम है तो दो मध्य मान होंगे mवाँ तथा (m+1)वाँ मूल्य। मध्यिका इन दोनों का समांतर माध्य होगा। हमने इनके उदाहरण पहले ही दे दिए हैं।

उदाहरण 10

निम्नलिखित आँकड़े किसी बस द्वारा दस बार में ले जाए गए सवारियाँ की संख्या दिखाते हैं :

22	26	14	30	18
11	35	41	12	32

जाँच करें कि ले जाए गए यात्रियों की संख्या का समांतर माध्य 24.1 है। मध्यिका संख्या जानने के लिए, हम संख्याओं को आरोही क्रम में लगाते हैं।

11	12	14	18	22	26	30	32	35	41

यहाँ दो मध्य मान हैं – 5वाँ और 6वाँ। मध्यिका इन दोनों का समांतर माध्य है, अर्थात्

$$\text{मध्यिका} = \frac{22+26}{2} = \frac{48}{2} = 24$$

वर्गीकृत आँकड़े

उदाहरण 11

मान लें कि x एक संतत चर है तथा मानों को वर्ग-अंतरालों मे वर्गीकृत किया गया है। इनका आवृत्ति वितरण नीचे सारणी में दिखाया गया है :

सारणी - 6.16

आवृत्ति वितरण

वर्ग-अंतराल x	आवृत्ति f
75-125	15
125-175	22
175-225	38
225-275	16
275-325	8
325-375	5
375-425	3
कुल	**107**

x के मूल्य का परिकलन करें।

मध्यिका का परिकलन करने के लिए, सबसे पहले ऊपर दी गई सारणी से संचयी आवृत्तियों का परिकलन करें।

सारणी - 6.17

संचयी आवृत्तियाँ

वर्ग-अंतराल	आवृत्ति	'से कम' संचयी आवृत्तियाँ
1	2	3
75-125	15	15=15
125-175	22	37=15+22
175-225	38	75=15+22+38
225-275	16	91=15+22+38+16
275-325	8	99=15+22+38+16+8
325-375	5	104=15+22+38+16+8+5
375-425	3	107=15+22+38+16+8+5+3
कुल	**107**	

x के मान आवृत्ति-सारणी में आरोही क्रम में लगाए गए हैं। हम देखते हैं कि 15 मान 125 से कम हैं; 15+22=37 मान 175 से कम हैं; 15+22+38 = 75 मान 225 से कम हैं, इत्यादि। इन्हें 'से कम' संचयी आवृत्तियाँ कहा जाता है।

चूँकि x एक संतत चर है, $\frac{N}{2}$वाँ मान $\left(\frac{107}{2} = 53.5\text{वाँ मान}\right)$ मध्यतम मान है, जो मध्यिका कहलाता है।

चूँकि 37 मान 175 से कम है तथा 75 मान 225 'से कम' है, $\frac{N}{2}$वाँ मान (53.5वाँ मान) वर्ग (175-225) के अंतर्गत है।

अतः (175-225) **मध्यिका वर्ग** है। हम मान लेते हैं कि इस वर्ग में सभी मान समान रूप से वितरित हैं।

अब, नीचे दिए गए सूत्र का प्रयोग मध्यिका (अर्थात् 53.5वाँ मान) को अन्तर्वेशित (interpolate) करने के लिए करें :

$$\text{मध्यिका} = L + \frac{\frac{N}{2} - c.f.}{f} \times h$$

जहाँ,

L = मध्यिका वर्ग की निचली सीमा (अर्थात् 175)

N = कुल आवृत्ति (अर्थात् 107)

c.f. = संचयी आवृत्ति = निम्नतम सीमा L से कम मानों की संख्या (अर्थात् 37)
f = मध्यिका वर्ग में आवृत्ति (अर्थात् 38)
h = वर्ग अंतराल की चौड़ाई (अर्थात् 50)

अतः उपर्युक्त सूत्र में प्रतिस्थापित करते हुए,

$$\text{मध्यिका} = 175 + \frac{53.5-37}{38} \times 50$$

$= 196.7$

इन मान की तुलना समांतर माध्य से करने के लिए यह जानना चाहिए कि इस स्थिति में समांतर माध्य 203.27 है।

13. बहुलक का परिकलन

बहुलक x का वह मान है जिसके लिए आवृत्ति सबसे अधिक है। उदाहरण के लिए, यदि x के मान हैं 70, 69, 69, 75, 75, 75, 60 तथा 80 तो स्पष्ट है कि बहुलक (या बहुलक मूल्य) 75 होगा।

यदि x के मूल्य वर्गों में वर्गीकृत किए जाएँ (जैसे वे किसी वर्ग में समान रूप से वितरित हों) तथा इसके लिए हमें एक आवृत्ति वितरण भी दिया जाए तो :

क. उस वर्ग की पहचान करें जिसकी आवृत्ति सबसे अधिक है तथा

ख. बहुलक का परिकलन इस प्रकार करें :

$$\text{बहुलक} = L + \frac{f_0 - f_1}{2f_0 - f_1 - f_2} \times h$$

जहाँ,
L = बहुलक वर्ग की निम्नतम सीमा
f_0 = उच्चतम आवृत्ति
f_1 = बहुलक वर्ग के पहले वाले वर्ग में आवृत्ति
f_2 = बहुलक वर्ग के बाद वाले वर्ग में आवृत्ति
h = वर्ग अंतराल की चौड़ाई।

देखें कि यदि $f_1 = f_2$ तो,

बहुलक $= L + \frac{1}{2}h$ अर्थात बहुलक, बहुलक-वर्ग के मध्यतम मूल्य के बराबर है।

उदाहरण 12

निम्नलिखित आवृत्ति वितरण में 352 स्कूल विद्यार्थियों की लंबाई के संबंध में आँकड़े दिए गए हैं। बहुलक लंबाई ज्ञात करें।

लंबाई (फुटों में)	विद्यार्थियों की संख्या
3.0-3.5	12
3.5-4.0	37
4.0-4.5	79
4.5-5.0	152
5.0-5.5	65
5.5-6.0	7
कुल	**352**

चूँकि 152 उच्चतम आवृत्ति है अतः बहुलक वर्ग (4.5-5.0) है। अब, L=45, f_0=152, f_1=79, f_2=65 तथा h=0.5, अतः,

$$\text{बहुलक} = 4.5 + \frac{152-79}{2^2\, 152-70-65}{}^2\, 05$$

4.73 फुट (लगभग)

14. समांतर माध्य, मध्यिका तथा बहुलक की सापेक्ष स्थिति

मान लें कि

समांतर माध्य $= M_e$
मध्यिका $= M_i$
बहुलक $= M_0$

अतः e, i तथा 0 प्रत्यय (suffixes) हैं। इन तीनों का सापेक्ष महत्त्व होता है :

$M_e > M_i > M_0$ या $M_e < M_i < M_0$

ये प्रत्यय वर्णों के क्रम के अनुसार हैं।

मध्यिका सदा समांतर माध्य तथा बहुलक के बीच में रहती है।

15. विभाजक मान—चतुर्थक (Partition Values—Quartiles)

मध्यिका चरों के पूरे समुच्चय को दो समान भागों में बाँटती है। इसी प्रकार, चतुर्थक चरों के पूरे

समुच्चय को चार समान भागों में बाँटते हैं। चतुर्थक तीन होते हैं Q_1 Q_2 तथा Q_3।

Q_1 को प्रथम या निम्न चतुर्थक कहते हैं। समुच्चय के कुल मूल्यों की संख्या में, Q_1 से कम मूल्यों की संख्या एक-चौथाई है तथा Q_1 से अधिक मूल्यों की संख्या तीन-चौथाई।

Q_3 को तृतीय या उच्च चतुर्थक कहते हैं। समुच्चयों के कुल मूल्यों की संख्या में, Q_3 से कम मूल्यों की संख्या तीन-चौथाई है तथा Q_3 से अधिक मूल्यों की संख्या एक-चौथाई।

Q_2 को मध्यिका या द्वितीय चतुर्थक कहते हैं। अतः Q_1 तथा Q_2 के बीच मूल्यों की संख्या समुच्चय के कुल मूल्यों की संख्या का एक चौथाई है। Q_2 तथा Q_3 के बीच की संख्या भी कुल मूल्यों की संख्या का एक चौथाई है।

Q_1 तथा Q_3 के बीच के मूल्यों की संख्या समुच्चय के कुल मूल्यों की संख्या का आधा हिस्सा है।

चतुर्थक का परिकलन करते समय हम यह मानते हैं कि चरों के मूल्य किसी भी वर्ग में समान रूप से वितरित हैं।

उदाहरण 13

निम्नलिखित आवृत्ति वितरण के लिए Q_1 तथा Q_3 चतुर्थक का परिकलन करें :

वर्ग अंतराल	आवृत्ति	संचयी आवृत्ति
0-25	23	23
25-50	64	87
50-75	115	202
75-100	51	253
100-125	30	283
125-150	17	300
कुल	**300**	

कुल आवृत्ति = N = 300, अतः $\frac{N}{4} = 75$

हम देखते हैं कि 23 मूल्य 25 से कम हैं तथा 87 मूल्य 50 से अधिक हैं। अतः, 75वाँ मूल्य (Q_1) वर्ग-अंतराल 25-50 में होना चाहिए।

Q_1 का परिकलन निम्नलिखित सूत्र द्वारा किया जाता है :

$$Q_1 = L + \frac{\frac{N}{4} - c.f.}{f} \times h$$

जहाँ,

L = उस वर्ग की निम्न सीमा है जिसमें Q_1 स्थित है।
N = कुल आवृत्ति।
c.f. = L से कम मूल्यों की संख्या।
f = उस वर्ग की आवृत्तियाँ जिसमें Q_1 स्थित है।
h = वर्ग अंतराल की चौड़ाई।

हमारे उदाहरण में

L=25, $\frac{N}{4}$ = 75, c.f. = 23, f = 23, f=64 तथा h=25। अतः इन मूल्यों को उपयुक्त सूत्र में प्रतिस्थापित करते हुए हमें, $Q_1 = 45.3125$ मिलता है।

Q_3 निकालने के लिए भी हम ऐसी ही विधि का प्रयोग करते हैं।

Q_3 $\left(\frac{3}{4}N\right)$ वाँ मूल्य है।

चूँकि N = 300, अतः $\left(\frac{3}{4}N\right)$ वाँ मूल्य = 225 होगा। अतः यहाँ हमें 225 वें मूल्य का पता लगाना है।

आवृत्ति वितरण के अनुसार, हम देखते हैं कि 202 मूल्य 75 से कम तथा 253 मूल्य 100 से कम हैं। 225वाँ मूल्य वर्ग 75-100 में होगा।

Q_3 का परिकलन भी उसी प्रकार के सूत्र के प्रयोग द्वारा किया जाता है जैसा Q_1 के परिकलन के लिए किया गया था :

$$Q_{3=} = L + \frac{\frac{3}{4}N - c.f.}{f} \times h$$

जहाँ,

L = उस वर्ग की निम्न सीमा है जिसमें Q_3 स्थित है।
N = कुल आवृत्ति।
c.f. = L से कम मूल्यों की संख्या।
f = उस वर्ग में आवृत्ति जिसमें Q_3 स्थित है।
h = वर्ग अंतराल का आकार।

ऊपर दिए गए उदाहरण में Q_3 वर्ग-अंतराल 75-100 में स्थित है।

यहाँ दिए गए मूल्य इस प्रकार हैं :

$=75, \frac{3}{4}N=225, c.f.=202, f_1=51$ तथा $h=25$

इन मूल्यों को ऊपर दिए गए सूत्र में प्रतिस्थापित करते हुए हमें $Q_3 = 86.2745$ मिलता है।

अभ्यास

1. केंद्रीय प्रवृत्ति के मापों के रूप में समांतर माध्य, मध्यिका तथा बहुलक की तुलना करें। उन स्थितियों का वर्णन करें जिनमें एक औसत दूसरे औसत से अधिक उपयुक्त हैं?
2. समांतर माध्य चरों के मूल्यों के वितरण का गुरुत्व केंद्र कहलाता है। वर्णन करें।
3. समांतर माध्य से चरों के मूल्यों के विचलनों का योग शून्य के बराबर होता है। स्पष्ट करें।
4. यदि $x_1, \ldots\ldots, x_n$ चर x के n मूल्य हैं तथा A कोई इच्छित मूल्य है, ताकि $\Sigma(x_i - A)$, A से विचलनों का योग है ; बताएँ कि $\Sigma(x_i - A) = n(\bar{x} - A)$, जहाँ x का समांतर माध्य $\bar{x} = \frac{N}{2}\Sigma x_i$ है। अतः $\Sigma(x_i - A) = 0$, यदि केवल $A = \bar{x}$।
5. 'समांतर माध्य बहुत बड़े तथा बहुत छोटे मूल्यों से प्रभावित होता है, परंतु मध्यिका तथा बहुलक इनसे प्रभावित नहीं होते हैं।' व्याख्या करें।
6. किसी इलाके में 20 परिवारों का सब्जियों पर दैनिक व्यय (रु. में) नीचे दिया गया है :

25.00	26.50	30.25	28.00	23.00
31.40	34.00	33.00	30.50	27.20
28.00	35.00	38.60	34.00	22.50
24.00	23.70	28.00	29.00	32.20

दिखाएँ कि मध्यिका का मूल्य समांतर माध्य तथा बहुलक के बीच में स्थित है।

7. निम्नलिखित आवृत्ति वितरण के लिए समांतर माध्य, मध्यिका तथा बहुलक का परिकलन करें।

वर्ग अंतराल	आवृत्ति
1-3	3
3-5	10
5-7	16
7-9	13
9-11	8
11-13	3
13-15	1
	54

इनकी सापेक्ष स्थितियों की जाँच करें। क्या मध्यिका समांतर माध्य तथा बहुलक के बीच में स्थित है?

8. 50 विनिर्माण फर्मों को 50 प्रश्नावलियाँ बाँटी गईं। इनसे केवल 35 वापस प्राप्त की गईं। 'क्षमता प्रयोग' के कॉलम में निम्नलिखित आँकड़े दिए गए थे —

54.2	96.8	74.2	57.8	99.7	84	81.2
94.3	93.7	61.6	100.00	110.1	68	71
95.2	98.3	84	91.9	84	94.1	73
84	52.6	49.2	45	62.6	75.1	84
69.2	90.7	85.4	87.3	66.1	73.9	89.3

क. समांतर माध्य, मध्यिका तथा बहुलक का परिकलन करें तथा उनकी सापेक्ष स्थितियों का परीक्षण करें।

ख. आँकड़ों को उपयुक्त वर्गों में वर्गीकृत करें तथा उसके बाद वर्गीकृत आँकड़ों से समांतर माध्य, मध्यिका तथा बहुलक का परिकलन करें। प्राप्त परिणाम भाग (क) में प्राप्त परिणामों से किस प्रकार भिन्न है?

ग. किसी परीक्षा में 100 परीक्षार्थी थे जिनमें, 6 को विशिष्टता मिली, 18 को द्वितीय श्रेणी मिली, 43 को तृतीय श्रेणी मिली तथा 21 अनुत्तीर्ण हो गए। यदि विशिष्टता के लिए 75 प्रतिशत अंकों की, प्रथम श्रेणी के लिए 60 प्रतिशत अंकों की, द्वितीय श्रेणी के लिए 50 प्रतिशत की तथा उत्तीर्ण होने के लिए 40 प्रतिशत की आवश्यकता है, तो अंकों के वितरण की मध्यिका का परिकलन करें।

9. निम्नलिखित आवृत्ति वितरण के लिए उच्च तथा निम्न चतुर्थक का परिकलन करें :

वर्ग अंतराल	**आवृत्ति**
13-25	6
25-37	11
37-49	23
49-61	7
61-73	3
कुल	**50**

10. प्रश्न 10 में दिए गए वितरण के लिए मध्यिका का परिकलन करें। ज्यामितीय विधि से दो तोरणों तथा उनके प्रतिच्छेद बिंदु के द्वारा मध्यिका को दिखाएँ।

11. 50 परिवारों के दैनिक व्यय (रु. में) नीचे दिए गए है :

दैनिक व्यय (रु. में)	**परिवारों की संख्या**
100-150	3
150-200	9
200-300	21
300-500	10
500-1000	5
1000 से ऊपर	2
कुल	**50**

(इस उदाहरण में अंतिम वर्ग विवृत है। अतः हम उस वर्ग का माध्य मान निर्धारित नहीं कर सकते। दिए गए आँकड़ों के लिए समांतर-माध्य का परिकलन नहीं किया जा सकता। हालाँकि हम किसी आधार पर एक उच्च वर्ग सीमा-निर्धारित कर सकते हैं तथा समांतर माध्य निकाल सकते हैं। परंतु यह केवल एक अनुमान ही हो सकता है। मध्यिका, बहुलक तथा चतुर्थकों के परिकलन में कोई समस्या नहीं होगी।)

क. समांतर माध्य, मध्यिका तथा बहुलक का परिकलन करें।

ख. उच्च तथा निम्न चतुर्थकों का परिकलन भी करें।

परिक्षेपण के माप

जहाँ केंद्रीय प्रवृत्ति के माप वितरण के केंद्रीय मूल्य की जानकारी देते हैं, वही परिक्षेपण (Dispersion) के माप की आवश्यकता केंद्रीय मूल्यों के आसपास मूल्यों के विचरण (परिक्षेपण या प्रकीर्ण) के परिमाण को मापने के लिए होती है। उदाहरण के लिए, मान लें कि पाँच परिवारों की मासिक आय (रु. में) निम्नलिखित है :

मासिक आय (रु. में)

4500	6000	5500	3750	4700

आय का समांतर माध्य 4890 रु. है तथा मध्यिका आय 4700 रु. है।

आय के विचरण का परिमाण केंद्रीय मूल्य से विचलनों द्वारा दिखाया जाता है।

समांतर माध्य से विचलन

-390	1110	610	-1140	-190

मध्यिका से विचलन

-200	1300	800	-950	0

यहाँ कुछ विचलन धनात्मक हैं तथा कुछ ऋणात्मक। कुछ विचलन अधिक हैं तथा कुछ कम। इसलिए केंद्रीय मूल्य से सभी मूल्यों के विचरणों के एक संक्षिप्त माप की आवश्यकता है। यह संक्षिप्त माप **परिक्षेपण की माप** कहलाती है।

इस अध्याय में हम परिक्षेपण के इन मापों का वर्णन करेंगे —

क. परास

ख. चतुर्थक विचलन

ग. माध्य विचलन

घ. मानक विचलन

माध्य विचलन तथा मानक विचलन को केंद्रीय मूल्य से विचलनों के रूप में परिभाषित कर सकते हैं। परास तथा चतुर्थक विचलन किसी विशेष मूल्य से विचलनों पर आधारित नहीं होते।

1. परास (Range)

दिए गए मूल्यों के किसी समुच्चय के सबसे बड़े तथा सबसे छोटे मूल्य के बीच के अंतर को परास कहते हैं। मान लें कि x के मूल्य आरोही क्रम में लगाए गए हैं, जैसे —

$x_1 < x_2 < \dots\dots\dots < x_n$.

ताकि यदि x_n सबसे बड़ा तथा x_1 सबसे छोटा मूल्य है, तो परास की परिभाषा इस प्रकार दी जा सकती है —

$R = x_n - x_1$

अतः x के सभी मूल्य परास के अंदर होते हैं। यदि परास का आकार बड़ा हो तो x के मूल्यों का फैलाव अधिक होगा जिससे x के मूल्यों का विचरण अधिक होगा। दूसरी ओर, यदि परास छोटा होगा तो मूल्यों का फैलाव कम होगा तथा x का विचरण भी कम होगा।

2. परास की विशेषताएँ

परास की निम्नलिखित विशेषताएँ हैं :

क. **यह सुपरिभाषित होता है।**

ख. **इसका परिकलन सरल है तथा इसे समझना आसान है।**

ग. **यह चर के प्रत्येक मूल्य पर निर्भर नहीं करता।** उच्चतम और निम्नतम मान के बीच मूल्यों के वितरण से परास पर कोई असर नहीं होता।

घ. **यह अंतिम मूल्यों द्वारा आवश्यकता से अधिक प्रभावित होता है।** उदाहरण के लिए, ऊपर दिए गए पाँच परिवारों की मासिक आय को लें। चूँकि सबसे अधिक आय 6000 रु. प्रति माह है तथा सबसे कम 3750 रु. प्रतिमाह है, इसलिए परास 2250 रु. होगा। यदि हम छठे परिवार की आय को लें, जो 15,000 रु. प्रति माह है, तो परास में पाँच गुणा वृद्धि हो जाती है। यह 11,250 रु. के बराबर हो जाता है (5 x 2250)।

ङ. **परास चर के मापन की इकाइयों पर निर्भर करता है।** इसके मापन की इकाई भी वही होती है जो चर की होती है। आय का परास रु. में दिखाते हैं, दूरी का परास किलोमीटर में, आदि।

3. चतुर्थक विचलन (Quartile Deviation)

चरों के दिए गए मूल्यों के लिए (वर्गीकृत या अवर्गीकृत आँकड़ों के लिए) हम उच्च चतुर्थक Q_3 तथा निम्न चतुर्थक Q_1 का परिकलन कर सकते हैं, जैसा पिछले अध्याय में बताया गया है। अतः **चतुर्थक विचलन** निम्नलिखित होगा :

$$\frac{Q_3 - Q_1}{2}$$

इसे **'अर्ध-अंतश्चतुर्थक परास'**(Semi-Inter Quartile Range) भी कहते हैं। वस्तुतः $Q_3 - Q_1$ 'अंतश्चतुर्थक परास' है।

चतुर्थक विचलन के प्रयोग का लाभ यह है यह अंतिम मूल्यों से प्रभावित नहीं होता, जबकि परास इनसे प्रभावित होता है। अन्य दृष्टियों से चतुर्थक विचलन की विशेषताएँ परास की विशेषताओं के समान हैं। ये विशेषताएँ नीचे दी गई हैं :

क. **यह सुपरिभाषित होता है।**

ख. **इसका परिकलन सरल है तथा इसे समझना आसान है।**

ग. **यह चर के प्रत्येक मूल्य पर निर्भर नहीं करता है।**

घ. **चतुर्थक विचलन के मापन की इकाई भी वही होती है जो चर के मापन की होती है।**

उदाहरण 1

किसी चर x का आवृत्ति वितरण नीचे दिया गया है:

वर्ग अंतराल x	आवृत्ति f
10-20	4
20-40	10
40-70	26
70-120	8
120-200	2
कुल	**50**

परास तथा चतुर्थक विचलन का परिकलन करें तथा विचलनों के दोनों मापों की तुलना करें।

यह स्पष्ट है कि परास 190 है क्योंकि अधिकतम मूल्य 200 है तथा न्यूनतम मूल्य 10 है।

चतुर्थक विचलन के परिकलन के लिए, सबसे पहले संचयी आवृत्तियों का परिकलन करें, जैसा नीचे सारणी में दिखाया गया है :

सारणी - 7.1

संचयी आवृत्तियों का परिकलन

वर्ग-अंतराल x	आवृत्ति f	संचयी आवृत्तियाँ
10-20	4	4
20-40	10	14
40-70	26	40
70-120	8	48
120-200	2	50
कुल	**50**	

Q_1 का परिकलन

परिभाषा के अनुसार Q_1 से कम मूल्यों की संख्या $\frac{1}{4} \times 50 = 125$ है। हम मान लेते हैं कि x के सभी मूल्य किसी वर्ग में समान रूप से वितरित हैं। चूँकि x के मूल्य आरोही क्रम में रखे गए हैं, उस वर्ग-अंतराल की पहचान करें जिसमें 12.5वाँ मूल्य हो।

वह वर्ग जिसमें 12.5वाँ मूल्य स्थित है 20-40 का वर्ग है, क्योंकि 4 मूल्य 20 से कम तथा 14 मूल्य 40 से कम है।

सूत्र का प्रयोग करते हुए,

$$L+\frac{\frac{N}{4}-c.f.}{f}\times h$$

जहाँ,

$L = 20, \frac{N}{4} = 12.5, c.f. = 4, f = 10$ तथा $h = 20$।

इन मूल्यों को सूत्र में प्रतिस्थापित करते हुए,

$$Q_1 = 20 + \frac{12.5-4}{10} \times 20 = 37$$

इसी प्रकार Q_3 निकालें जो $\frac{3}{4}\times 50 = 37.5$ वाँ मूल्य है। जिस वर्ग में Q_3 है वह 40–70 का वर्ग है, क्योंकि 14 मूल्य 40 से कम तथा 40 मूल्य 70 से कम हैं। अतः पहले दिए गए सूत्र का प्रयोग करते हुए :

$$Q_3 = 40 + \frac{37.5-14}{26} \times 30 = 67.11$$

इसलिए, चतुर्थक विचलन $\frac{Q_3-Q_1}{2} = \frac{67.11-37}{2}$

$= 15.06 \approx 15.1$ है।

इसके मापन की इकाई भी वही है जो x के मापन की इकाई है।

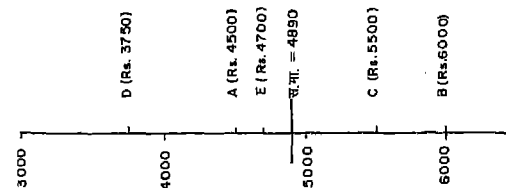
4. केंद्रीय मूल्य से विचलनों के रूप में परिक्षेपण के माप

हम पहले देख चुके हैं कि परास तथा चतुर्थक विचलन परिक्षेपण के सामान्य माप हैं। ये वितरण के किसी विशेष मूल्य की ओर संकेत नहीं करते। परंतु, केंद्रीय मूल्य या किसी अन्य मूल्य से विचलन उस मूल्य से परिक्षेपण की एक बेहतर जानकारी देते हैं। उदाहरण के लिए, पाँच परिवारों की मासिक आय को देखें जो नीचे दी जा रही है :

परिवारों की मासिक आय

परिवार	क	ख	ग	घ	ङ.
आय (रु.में)	4500	6000	5500	3750	4700

आय का समांतर माध्य 4890 रु. है तथा 4890 रु. से आय के विचलन नीचे दिए गए हैं :

-390	1110	610	-1140	-190

आरेख 7.1 : समांतर माध्य से विचलन

जैसा आरेख 7.1 में दिखाया गया है, परिवार घ समांतर माध्य की बायीं ओर सबसे अधिक दूरी पर है तथा परिवार ख इसकी दायीं ओर सबसे अधिक दूरी पर। अन्य परिवारों की आय दोनों ओर से समांतर माध्य के करीब है। इसी प्रकार, हम किसी अन्य मूल्य का चुनाव कर उससे विचलन दिखा सकते हैं।

सामान्य तौर पर, कुछ विचलन धनात्मक होते हैं और कुछ ऋणात्मक। कुछ बड़े होते हैं और कुछ छोटे।

परिक्षेपण की पूरी जानकारी के लिए, हम इन विचलनों के औसतों को ले सकते हैं। परंतु विचलनों का योग करते समय, हम देखते हैं कि धनात्मक तथा ऋणात्मक विचलन एक-दूसरे को निष्क्रिय कर देतें हैं तथा उनका योग उपक्षेणीय होता है। वस्तुतः समांतर माध्य से विचलनों का योग हमेशा शून्य होता है। अतः विचलनों का सीधा योग करने से कोई लाभ नहीं। इसके स्थान पर हम 'निरपेक्ष मूल्यों' (उनके चिह्नों की उपेक्षा करते हुए) या 'वर्गीकृत विचलनों' को ले सकते हैं।

निम्नलिखित अनुच्छेदों में हम परिक्षेपण के मापों को समांतर माध्य से चरों के मूल्यों के **(क) निरपेक्ष विचलन** तथा **(ख) विचलनों के वर्गों** के रूप में परिभाषित करेंगे।

5. माध्य विचलन (Mean Deviation)

मान लें कि $x_1, x_2, \ldots\ldots x_n$ आदि x के n मूल्य हैं तथा $\bar{x} = \frac{1}{n}\Sigma x_i$ उनका समांतर माध्य है तथा $\bar{x}$ से उनके विचलन निम्नलिखित हैं :

$$x_1 - \bar{x}, x_2 - \bar{x}, \ldots\ldots\ldots, x_n - \bar{x}$$

इनमें से कुछ विचलन धनात्मक होंगे तथा कुछ ऋणात्मक। विचलनों के चिह्नों की उपेक्षा करते हुए :

$$|x_1 - \bar{x}|, \; |x_2 - \bar{x}|, \ldots\ldots\ldots, |x_n - \bar{x}|$$

ये विचलनों के निरपेक्ष मूल्य हैं, जहाँ दो समांतर रेखाएँ || यह दर्शाती हैं कि केवल निरपेक्ष मूल्य लिए गए हैं। [इसे हम माड्यूलस मूल्य (Modulus Value) कहते हैं।]

निरपेक्ष विचलनों के समांतर माध्य को **माध्य विचलन** या **माध्य निरपेक्ष विचलन** कहते हैं। अतः

$$\frac{1}{n}\Sigma|x_i - \bar{x}|$$

जो समांतर माध्य से x का माध्य-विचलन है।

उदाहरण 2

निम्नलिखित सारणी में राजस्थान के ऐसे 11 प्रमुख जिलों की जनसंख्या का अनुमान दिया गया है जहाँ की जनसंख्या 20 लाख से अधिक है। जनसंख्या के परास, चतुर्थक विचलन तथा माध्य विचलन का परिकलन करें।

चूँकि, जयपुर की जनसंख्या सबसे अधिक 49,21,000 है तथा अजमेर की सबसे कम 20,13,000। अतः परास 29,08,000 है। मापन की इकाई 'व्यक्तियों की संख्या' है।

सारणी - 7.2

राजस्थान के प्रमुख जिलों* की जनसंख्या का अनुमान (2001)

जिले	**जनसंख्या ('000)**
अजमेर	2013
अलवर	2691
भरतपुर	2932
गंगानगर	2937
जयपुर	4921
जोधपुर	2275
कोटा	2234
नागौर	2606
सवाई माधोपुर	2315
सीकर	2019
उदयपुर	3389

* प्रमुख जिले वे हैं जिनकी जनसंख्या 20 लाख से अधिक है।

चतुर्थक के परिकलन के लिए, हम आँकड़ों को जनसंख्या के आकार के अनुसार आरोही क्रम में लगाते हैं :

2013, 2019, $\boxed{2234}$, 2275, 2315, $\boxed{2606}$, 2691, 2932, $\boxed{2937}$, 3389, 4921

विभाजक मूल्य हैं $Q_1 = 2234$, $Q_2 = 2606$, तथा $Q_3 = 2937$, जो $\frac{1}{4}(n+1)$वाँ, $\frac{2}{4}(n+1)$ तथा $\frac{3}{4}(n+1)$वाँ मूल्य हैं, जहाँ $n = 11$

इसलिए चतुर्थक विचलन निम्नलिखित होगा :

$$\frac{Q_3 - Q_1}{2} = 351.5$$

मापन की इकाई ' व्यक्तियों की संख्या' है।

जनसंख्या का समांतर माध्य है

$$\frac{30332}{11} \approx 2754.45$$

सारणी - 7.3 में समांतर माध्य से विचलन दिए गए हैं :

सारणी - 7.3

समांतर माध्य से विचलन

जिले	**जनसंख्या ('000) x**	**समांतर माध्य = 2757.45 से विचलन $(x-\bar{x})$**	**मध्यिका = 2606 से विचलन $(x-M_e)$**
अजमेर	2013	-744.45	-593
अलवर	2691	-66.45	85
भरतपुर	2932	174.55	326
गंगानगर	2937	179.55	331
जयपुर	4921	2163.55	2315
जोधपुर	2275	-482.45	-331
कोटा	2234	-523.45	-372
नागौर	2606	-151.45	0
सवाई माधोपुर	2315	-442.45	-291
सीकर	2019	-738.45	-587
उदयपुर	3389	631.55	783

इसका पुनर्परीक्षण किया जाना चाहिए कि समांतर माध्य से विचलन शून्य है। परंतु विचलनों के चिह्नों की उपेक्षा करते हुए :

$$\sum \left(x_i - \bar{x}\right) = 6298.35$$

तथा माध्य से माध्य विचलन $= \frac{1}{N} \sum \left|x_i - \bar{x}\right| = \frac{6298.35}{11} = 572.58$ (लगभग) है। मापन की इकाई 'व्यक्तियों की संख्या' है। यहाँ हमने माध्य 2606 निकाला है।

अब हम मध्यिका से माध्य विचलन पर विचार करें। मध्यिका से माध्य विचलन निम्नलिखित होगा :

$$\frac{1}{n} \sum |x_i - \text{मध्यिका}| = \frac{6014}{11} = 546{\cdot}73$$

मापन की इकाई 'व्यक्तियों की संख्या' है।

यहाँ ध्यान दें की मध्यिका से माध्य विचलन समांतर माध्य से माध्य विचलन की तुलना में कम है। अतः सामान्यतः

माध्य विचलन जब मध्यिका से लिया जाता है तो यह न्यूनतम होता है।

विद्यार्थी अन्य मूल्यों से माध्य विचलनों के द्वारा इसकी जाँच कर सकते हैं।

6. **वर्गीकृत आँकड़ों से माध्य विचलन का परिकलन**

उदाहरण 3

किसी सौंदर्य प्रतियोगिता में 150 महिलाओं के कद के वितरण निम्नलिखित हैं :

कद (इंचों में)	**आवृत्ति**
62.0-63.5	12
63.5-65.0	20
65.0-66.5	28
66.5-68.0	18
68.0-69.5	19
69.5-71.0	20
71.0-72.5	30
72.5-74.0	3
कुल	**150**

समांतर माध्य तथा मध्यिका से परास, चतुर्थक विचलन तथा माध्य विचलन का परिकलन करें।

चूँकि सबसे अधिक कद 74" है तथा सबसे कम 62", इसलिए परास 12" है। चतुर्थकों के आकलन के लिए संचयी आवृत्तियाँ निम्नलिखित सारणी में दी गई हैं :

सारणी - 7.4

संचयी आवृत्तियाँ

कद	आवृत्ति	संचयी अवृत्तियाँ
62.0-63.5	12	12
63.5-65.0	20	32
65.0-66.5	28	60
66.5-68.0	18	78
68.0-69.5	19	97
69.5-71.0	20	117
71.0-72.5	30	147
72.5-74.0	3	150
कुल		**150**

$Q_1\left(\frac{1}{4}N\right)$ वाँ मूल्य है अर्थात् 37.5 वाँ मूल्य। यह 65.0-66.5 वें वर्ग में है। Q_1 को जानने के लिए सूत्र का प्रयोग करें :

$$Q_1 = 65 + \frac{37.5-32}{28} \times 1{\cdot}5 \approx 65{\cdot}29$$

Q_2 = मध्यिका $\frac{N}{2}$ वाँ मूल्य है अर्थात् 75 वाँ मूल्य। यह वर्ग Q_2 = 66.5-68.0 में है। Q_2 को प्राप्त करने के लिए सूत्र का प्रयोग करें।

$$\text{मध्यिका } Q_2 = 66.5 + \frac{75-60}{18} \times 1.5 = 67.75$$

$Q_3\left(\frac{3}{4}N\right)$ वाँ मूल्य है, अर्थात् 112.5 वाँ मूल्य। यह वर्ग 69.5-71.0 में है। अतः सूत्र का प्रयोग करते हुए :

$$Q_3 = 69.5 + \frac{112.5-97}{20} \times 1{\cdot}5 = 70.66$$

अतः चतुर्थक-विचलन होगा -

$$\frac{Q_3 - Q_1}{2} = \frac{70.66 - 65.29}{2} = 2.686$$

मापन की इकाइयाँ 'इंचों' में हैं।

कदों के समांतर माध्य का परिकलन सारणी 7.5 में दिखाया गया है। अतः

$$\bar{d'} = \frac{\sum f_i d_i'}{N} = \frac{\sum f_i d_i'}{\sum f} = \frac{207}{150} = 1.38$$

$$\bar{x} = 65.75 + 1.5 \text{x} 1.38 = 67.82$$

सारणी - 7.5

समांतर माध्य का परिकलन

(कल्पित माध्य 65.75 तथा समापवर्तक =1.5)

कद वर्ग-अंतराल	मध्य मान x	d = x – 65.76	d = d/1.5	आवृत्ति f	f.d'
62.0-63.5	62.75	-3	2	12	-24
63.5-65.0	64.25	-1.5	-1	20	-20
65.0-66.5	65.75	0	0	28	0
66.5-68.0	67.25	1.5	1	18	18
68.0-69.5	68.75	3	2	19	38
69.5-71.0	70.25	4.5	3	20	60
71.0-72.5	71.75	6	4	30	120
72.5-74.0	73.25	7.5	5	3	15
				$\sum f$ = 150	**207**

हम माध्य विचलनों का परिकलन **प्रत्यक्ष विधि** द्वारा इस प्रकार कर सकते हैं :

$$\frac{1}{N}\Sigma f_i|x_i-\bar{x}|;$$

या **पद-विचलन विधि** द्वारा इस प्रकार कर सकते हैं :

$$h\times\frac{1}{N}\Sigma f_i\left|d_i^{'}-\bar{d^{'}}\right|$$

सारणी - 7.6

माध्य विचलन का परिकलन
प्रत्यक्ष विधि द्वारा

($\bar{x}$ = 67.82)

मध्य मान x	$\|(x-\bar{x})\|$	f	$f\|(x-\bar{x})\|$
62.75	5.07	12	60.84
64.25	3.57	20	71.40
65.75	2.07	28	57.96
67.25	0.57	18	10.26
68.75	0.93	19	17.67
70.25	2.43	20	48.60
71.75	3.93	30	117.90
73.25	5.43	3	16.29
कुल		**Σ_f = 150**	**400.92**

माध्य विचलन $=\frac{1}{N}\Sigma f_i|x_i-\bar{x}|=\frac{400.92}{150}=2.6728$

सारणी - 7.7

माध्य विचलन का परिकलन
(पद-विचलन विधि द्वारा)
(कल्पित माध्य = 65.75 तथा समापवर्तक = 1.5)

$d^{'}=\frac{x-65.75}{1.5}$	$d^{'}-\bar{d^{'}}=d^{'}-1.38$	$\|d^{'}-\bar{d^{'}}\|$	f	$f\|d^{'}-\bar{d^{'}}\|$
-2	-3.38	3.38	12	40.56
-1	-2.38	2.38	20	47.60
0	-1.38	1.38	28	38.64
1	-0.38	0.38	18	6.84
2	0.62	0.62	19	11.78
3	1.62	1.62	20	32.40
4	2.62	2.62	30	78.60
5	3.62	3.62	3	10.86
कुल			**Σ_f = 150**	**267.28**

समांतर माध्य से माध्य विचलन

$$=\frac{1}{N}\Sigma f_i|x_i-\bar{x}|=h\times\frac{1}{N}\Sigma f_i\left|d_i^{'}-\bar{d^{'}}\right|$$

$$=15 \times \frac{267.28}{150}=2.6728$$

माध्य विचलन की निम्नलिखित विशेषताएँ हैं :

क. यह सुपरिभाषित होता है।

ख. यह चर के सभी मूल्यों पर निर्भर करता है।

ग. यह केंद्रीय मूल्य से निरपेक्ष विचलनों पर आधारित होता है।

घ. इसे समझना आसान है।

ङ. इसका परिकलन परास तथा चतुर्थक विचलन की तुलना में कठिन होता है।

च. यह बीजगणितीय विवेचन के योग्य होता है।

छ. माध्य-विचलन के मापन की इकाइयाँ चर x के मापन की इकाइयों के समान होती हैं।

7. मानक विचलन (Standard Deviation)

समांतर माध्य से निरपेक्ष विचलनों के स्थान पर हम प्रत्येक विचलनों का वर्ग निकालकर, वर्गीकृत विचलनों का समांतर माध्य निकालते हैं। इससे हमें मूल्यों का **विचरण मापांक** (Variance) प्राप्त होता है।

विचरण मापांक का **धनात्मक** वर्गमूल दिए गए मूल्यों का **मानक विचलन** कहलाता है।

मान लें कि $x_1, x_2, \ldots\ldots, x_n$ आदि x के n मूल्य हैं तो उनका समांतर माध्य

$$\bar{x}=\frac{1}{n}\Sigma x_i$$

तथा $x_1-\bar{x}$, $x_2-\bar{x}\ldots\ldots, x_n-\bar{x}$ आदि $\bar{x}$ से x के विचलन हैं। इसलिए, x का विचरण मापांक निम्नलिखित होगा :

$$\sigma^2=\frac{1}{n}\Sigma(x_i-\bar{x})^2$$

यह दिखाया जा सकता है कि

$$\sigma^2 = \frac{1}{n}\sum(x_i - \bar{x})^2 = \frac{1}{n}\sum x_i^2 - \bar{x}^2$$

जो परिकलनों के लिए अधिक सुविधाजनक है।

परंपरानुसार, विचरण मापांक को संकेत σ^2 द्वारा दिखाया जाता है, जहाँ σ एक ग्रीक अक्षर 'सिग्मा' है। वस्तुतः σ छोटा सिग्मा है तथा $\sum$ बड़ा सिग्मा। σ का वर्ग σ^2 है।

विचरण मापांक का वर्गमूल मानक विचलन है जो निम्नलिखित है :

$$\sigma = +\sqrt{\frac{1}{N}\sum\left(x_i - \bar{x}\right)^2}$$

हम वर्गमूल के पहले हमेशा (+) चिह्न का प्रयोग करते हैं। अतः σ हमेशा धनात्मक होता है।

विचरण मापांक के मापन की इकाइयाँ $_x{}^2$ की इकाइयों के रूप में तथा σ के मापन की इकाइयाँ 'x की इकाइयों' के रूप में दिखाई जाती हैं।

उदाहरण 4

सारणी 7.2 में राजस्थान के 11 प्रमुख जिलों में (जिनकी जनसंख्या 20 लाख से अधिक है) जनसंख्या का अनुमान (हज़ार में) दिया गया है। जनसंख्या के विचरण मापांक का परिकलन करें।

परिकलनों को सरल बनाने के लिए हम जनसंख्या के अनुमान को दस लाख के निकटतम तक लेते हैं।

प्रत्यक्ष विधि

इसके लिए, नीचे दिए गए सूत्र का प्रयोग करें :

$$\sigma_x^2 = \frac{1}{n}\sum x_i^2 - \bar{x}^2$$

जहाँ σ^2 के साथ x प्रत्यय का प्रयोग x के मूल्यों (अर्थात् जनसंख्या) के विचरण को दिखाता है। जिलों की जनसंख्या (लाख में) नीचे दी गई है

20, 27, 29, 29, 49, 23, 22, 26, 23, 20 तथा 34

इससे हमें निम्नलिखित प्राप्त होता है :

$$\bar{x} = \frac{302}{11} \approx 27.45 \text{ तथा } \frac{\sum x_i^2}{n} \approx 816.91$$

अतः

$$\sigma_x^2 \approx 63.4075 \text{ तथा } \sigma_x \approx 7.96$$

कल्पित माध्य विधि

विचरण मापांक कल्पित माध्य के चुनाव से प्रभावित नहीं होता।

मान लें कि कल्पित माध्य 20 है। अतः विचलन ($d_i = x_i - 20$) निम्नलिखित हैं :

0, 7, 9, 9, 29, 3, 2, 6, 3, 0 तथा 14

इससे हमें निम्नलिखित प्राप्त होगा :

$$\bar{d} = \frac{\sum d_i}{n} = \frac{82}{11} \approx 7.45$$

तथा

$$\frac{\sum d_i^2}{n} = \frac{1306}{11} \approx 118.73$$

अतः, d_i का विचरण मापांक निम्नलिखित होगा

$$\sigma_d^2 = \frac{\sum d_i^2}{n} - \bar{d}^2 \approx 63.23$$

यह ऊपर प्राप्त किए गए x के विचरण मापांक (σ_x^2) के समान है (त्रुटि को पूर्णांक बनाने को छोड़कर)। x का मानक-विचलन d की भाँति है :

$$\sigma_x = +\sqrt{63.23} \approx 63.23$$

पद-विचलन विधि

पहले एक सुविधाजनक समापवर्तक h चुनें तथा कल्पित माध्य से प्रत्येक विचलन को h से भाग दें तथा $d'_i = \frac{d_i}{h}$ निकालें जहाँ i=1,........,n ।

तब

$$\bar{d'} = \frac{\sum d'_i}{n}$$

तथा $d'_1, \ldots, d'_n$ का विचरण मापांक निम्नलिखित होगा :

$$\sigma_{d'}^2 = \frac{\sum d_i'^2}{n} - (\bar{d'_1})^2$$

x का विचरण मापांक :

$$\sigma_x^2 = h^2\sigma_d^2$$

ताकि

$$\sigma_x = h\sigma_d.$$

इस उदाहरण में, पद-विचलन विधि के प्रयोग का कोई लाभ नहीं क्योंकि कल्पित माध्य से विचलनों के आकार छोटे हैं।

8. वर्गीकृत आँकड़ों से मानक विचलन का परिकलन

हम वर्गीकृत आँकड़ों से मानक विचलन के परिकलन के लिए उदाहरण 3 में दिए गए आवृत्ति वितरण का प्रयोग करते हैं।

उदाहरण 5

किसी सौंदर्य प्रतियोगिता में 150 महिलाओं के कदों का आवृत्ति वितरण नीचे दिया गया है :

कद (इंचों में)	आवृत्ति
62.0-63.5	12
63.5-65.0	20
65.0-66.5	28
66.5-68.0	18
68.0-69.5	19
69.5-71.0	20
71.0-72.5	30
72.5-74.0	3
कुल	**150**

कदों के मानक विचलन का परिकलन करें।

प्रत्यक्ष विधि

मानक विचलनों का परिकलन नीचे सारणी 7.8 में दिखाया गया है।

सारणी - 7.8

मानक विचलनों का परिकलन

(प्रत्यक्ष विधि द्वारा)

कद (इंचों में)	मध्य मान x	आवृत्ति f	f.x	$f.x^2$
62.0-63.5	62.75	12	753.00	47250.75
63.5-65.0	64.25	20	1285.00	82561.25
65.0-66.5	65.75	28	1841.00	121045.75
66.5-68.0	67.25	18	1210.50	81406.125
68.0-69.5	68.75	19	1306.25	89804.6875
69.5-71.0	70.25	20	1405.00	98701.25
71.5-72.0	71.75	30	2152.50	15441.875
72.5-74.0	73.25	3	219.75	16096.6875
कुल		**150**	**10173**	**69138.375**

इसलिए,

$$\bar{d} = \frac{\Sigma f_i x_i}{N} = \frac{10173}{150} = 67.82$$

तथा

$$\frac{\Sigma f_i x_i^2}{N} = 4608.7225$$

विचरण के परिकलन के चरण नीचे दिए जा रहे हैं :

चरण 1 : वर्ग अंतरालों (कॉलम 2) के मध्य मान (x_i) निकालें।

चरण 2 : कॉलम 3 की आवृत्तियों (f_i) को x_i के संगत मूल्यों (कॉलम 2) से गुणा करें जिससे $f_i x_i$ (कॉलम 4) प्राप्त होता है।

चरण 3 : कॉलम (4) का योग निकालें जिससे $\Sigma f_i x_i$ मिलें, तथा

$$\bar{x} = \frac{\Sigma f_i x_i}{N} = N = \Sigma f_i = 150$$

चरण 4 : $f_i x_i$ (कॉलम 4) को संगत x_i (कॉलम 2) से गुणा करें तथा $f_i x_i^2$ (कॉलम 5) निकालें।

चरण 5 : कॉलम (5) का योग निकालें जिससे

$\Sigma f_i x_i^2$ तथा $\frac{\Sigma f_i x_i^2}{N}$ मिले।

चरण 6 : $\sigma_x^2 = \frac{\Sigma f_i x_i^2}{N} - \overline{x^2}$ निकालें।

कदों (x_i) के विचरण मापांक की इकाइयों x के मापन की इकाइयों के वर्ग के रूप में दिखाई गई हैं तथा x के मानक विचलन के मापन की इकाइयाँ x के मापन की इकाइयों के समान हैं।

कल्पित माध्य विधि

निम्नलिखित सारणी में कल्पित माध्य विधि द्वारा मानक विचलन का परिकलन दिखाया गया है :

सारणी - 7.9

मानक विचलन का परिकलन
(कल्पित माध्य विधि द्वारा)
(कल्पित माध्य = 67.25)

मध्य मान x	**d = x − 67.25**	**f**	**f.d**	**f.d²**
1	2	3	4	5
62.75	- 4.5	12	-54	243
64.25	-3.0	20	-60	180
65.75	-1.5	28	-42	63
67.25	0	18	0	0
68.75	1.5	19	28.5	42.75
70.25	3.0	20	60	180
71.75	4.5	30	135	607.5
73.25	6.0	3	18	108
		150	**85.5**	**1424.25**

इसलिए,

$$\bar{d} = \frac{\Sigma f_i x_i}{N} = \frac{85.5}{150} = 0.57$$

$$\frac{\Sigma f_i d_i^2}{N} = \frac{1424.25}{150} = 9.495$$

$$\sigma_d^2 = \frac{\Sigma f_i d_i^2}{N} - \overline{d^2} = 9.1701\ ('')^2$$

यहाँ ध्यान दिया जाना चाहिए कि

$$\sigma_x^2 = \sigma_d^2 = 9.1701$$

अर्थात् **x के विचरण मापांक मूल-बिंदु के चुनाव से प्रभावित नहीं होते।**

x का मानक विचलन विचरण मापांक का धनात्मक वर्ग मूल होता है, अर्थात् $\sigma_x \approx 3.03''$

x के विचरण मापांक के आकलन में निम्नलिखित चरणों पर ध्यान दिया जाना चाहिए :

चरण 1 : कल्पित माध्य स्वेच्छा से 67.25 चुना गया है।

चरण 2 : $d_i = x_i - 67.2$ (कॉलम 2 में) निकालें जहाँ x_i विभिन्न वर्गों के मध्य मान हैं।

चरण 3 : $f_i d_i$ (कॉलम 4 में) निकालें तथा $\Sigma f_i d_i$ निकाल कर $\bar{d} = \frac{\Sigma f_i d_i}{N}$ निकालें।

चरण 4 : कॉलम (2) तथा कॉलम (4) के संगत मूल्यों को गुणा करके $f_i d_i$ (कॉलम 5 में) निकालें तथा $\Sigma f_i d_i^2$ निकालें।

चरण 5 : d_i के विचरण मापांक के रूप में $\sigma_d^2 = \frac{\Sigma f_i d_i^2}{N} = \overline{d^2}$ निकालें।

चरण 6 : d_i के विचरण मापांक x_i के विचरण मापांक के बराबर हैं।

चरण 7 : x का मानक विचलन x के विचरण मापांक के धनात्मक वर्गमूल के बराबर है।

पद-विचलन विधि

पद-विचलन विधि द्वारा मानक विचलन का परिकलन नीचे दिया जा रहा है :

सारणी - 7.10

मानक विचलन का परिकलन

(पद विचलन विधि द्वारा)

(कल्पित माध्य = 67.25 तथा समापवर्तक = 1.5)

मध्य मान x	d= x–67.25	d'=d/1.5	f	f.d'	f.d'^{2}
1	2	3	4	5	6
62.75	–4.5	–3	12	–36	108
64.25	–3.0	–2	20	–40	80
65.75	–1.5	–1	28	–28	28
67.25	0	0	18	0	0
68.75	1.5	1	19	19	19
70.25	3.0	2	20	40	80
71.75	4.5	3	30	90	270
73.25	6.0	4	3	12	48
कुल			**150**	**57**	**633**

इससे हमें निम्नलिखित प्राप्त होता है :

$$\bar{d'} = \frac{\Sigma f_i d_i'}{N} = \frac{57}{150} = 0.38$$

$$\frac{\Sigma f_i d_i^2}{N} = \frac{633}{150} = 4.22$$

इसलिए,

$$\sigma_{d'}^2 = \frac{\Sigma f_i d_i^2}{N} - \overline{d^2} = 4.0756$$

तथा

$$\sigma_x^2 = h^2 \sigma_{d'}^2 = 1.5^2 \times 4.0756 = 9.1701('')^2$$

x का मानक विचलन x के विचरण मापांक का धनात्मक वर्ग है, इसलिए $\sigma_x \approx 3.03''$

मानक विचलन के परिकलन के चरण नीचे दिए गए हैं :

चरण 1 : हम 67.25 को कल्पित माध्य तथा 15 को समापवर्तक मानते हैं।

चरण 2 : $d_i = x_i - 67.25$ (कॉलम 2 में) तथा $d_i' = \frac{d_i}{15}$ (कॉलम 3 में) निकालते हैं जहाँ x_i वर्गों के मध्य-मान हैं जो कॉलम 1 में दिखाए गए हैं।

चरण 3 : आवृत्तियों (f_i) को d_i के संगत मूल्यों से गुणा कर $f_i d_i'$ (कॉलम 5 में) निकालें।

चरण 4 : $\bar{d'} = \frac{\Sigma f_i d_i'}{N}$ निकालें

चरण 5 : कॉलम (3) तथा कॉलम (5) के मूल्यों को गुणा करके $f_i d_i'^2$ (कॉलम 6 में) निकालें।

चरण 6 : $\frac{\Sigma f_i d_i'^2}{N}$ तथा d' के विचरण मापांक निकालें :

$$\sigma_{d'}^2 = \frac{\Sigma f_i d_i'^2}{N} - \overline{d'}^2$$

चरण 7 : हम देखते हैं कि x का विचरण मापांक निम्न है :

$$\sigma_x^2 = h^2 \sigma_{d'}^2$$

तथा विचरण मापांक का धनात्मक वर्गमूल σ_x का मानक विचलन है।

पद-विचलन विधि के प्रयोग से परिकलन बहुत सरल हो जाते हैं।

9. परिक्षेपण के विभिन्न मापकों की तुलना

अब हम ऊपर वर्णित परिक्षेपण के चारों मापकों की विशेषताओं की तुलना करें।

क. ये सुपरिभाषित होते हैं

परास, चतुर्थक विचलन, माध्य विचलन तथा मानक विचलन ये चारों मापक सुपरिभाषित होते हैं। उनकी परिभाषाओं में कोई अस्पष्टता नहीं होती।

ख. इनका परिकलन आसान होता है

परास का परिकलन सबसे आसान है। चतुर्थक विचलन में उच्च और निम्न चतुर्थकों के परिकलनों की आवश्यकता होती है पर वे भी काफी सरल हैं। हालाँकि, माध्य विचलन तथा मानक विचलन में

कुछ सुव्यवस्थित परिकलनों की आवश्यकता पड़ती है, परंतु ये भी सरल ही हैं।

ग. इन्हें समझना आसान है

इन सभी मापकों को समझना आसान है। परास तथा चतुर्थक विचलन परिक्षेपण के सामान्य मापक हैं, माध्य विचलन तथा मानक विचलन परिक्षेपण का मापन किसी केंद्रीय मूल्य से विचलनों के रूप में करते हैं। अतः माध्य विचलन तथा मानक विचलन परास के अंतर्गत मूल्यों के परिक्षेपण की एक बेहतर जानकारी देते हैं।

घ. ये सभी मूल्यों पर आधारित होते हैं

परास तथा चतुर्थक विचलन सभी मूल्यों पर आधारित नहीं होते हैं, जबकि माध्य विचलन तथा मानक विचलन चरों के सभी विचलन तथा मानक विचलन चरों के सभी मूल्यों पर आधारित होते हैं। परास सबसे अधिक अंतिम मूल्यों द्वारा प्रभावित होता है।

ङ. ये बीजगणितीय विवेचन के योग्य होते हैं

विश्लेषण की दृष्टि से मानक विचलन सबसे आसान होता है। अन्य मापकों का विश्लेषण भी किया जा सकता है पर वे थोड़ा कठिन होते हैं।

10. परिक्षेपण के सापेक्ष मापक — विचरण के गुणांक

परिक्षेपण के ऊपर दिए गए सभी मापक निरपेक्ष मापक हैं। इन सभी के मापन की इकाइयाँ वही होती हैं जो चरों के मापन की। अतः यदि हम आय के विचरण (रुपये में) पर विचार कर रहे हों तो परास, चतुर्थक-विचलन, माध्य विचलन तथा मानक विचलन सभी रुपये में होंगे। यदि कद से.मी. में मापा जा रहा हो तो ये सभी मापक से.मी. में ही होंगे। परिक्षेपण के मापन की यह विशेषता तब कठिनाई उत्पन्न कर सकती है, जब हम मूल्यों के दो ऐसे समुच्चयों के परिक्षेपण की तुलना करना चाहते हों जिनके —

क. केंद्रीय मूल्य अलग-अलग हों, तथा/या

ख. मापन की इकाइयाँ अलग-अलग हों।

उदाहरण के लिए, मान लें कि हम आय का मापन रुपयों के स्थान पर पैसों में करना चाहते हैं। इससे परिक्षेपण में सौगुनी वृद्धि हो जाएगी, जिससे ऐसा लग सकता है कि आय का परिक्षेपण बढ़ गया है जबकि वस्तुतः इसमें कोई परिवर्तन नहीं होता।

मूल्यों के दो समुच्चयों के परिक्षेपण की तुलना भी कठिन होती है, यदि मापन की इकाइयाँ समान नहीं हों तो। उदाहरण के लिए भारत तथा संयुक्त राष्ट्र अमेरिका की आय के परिक्षेपण की तुलना करना कठिन होगा। अमेरिका की आय डॉलरों में तथा भारत की रुपयों में होगी।

इस कठिनाई को दूर करने के लिए, मापन की इकाई की समस्या का हल निकालना होगा। यह तभी हो सकता है जब हम **परिक्षेपण के सापेक्ष मापक** का प्रयोग करें जो एक **शुद्ध संख्या** हो (जो मापन की इकाइयों पर निर्भर नहीं करती हो) परिक्षेपण का सापेक्ष मापक विचरण का गुणांक कहलाता है। इसे अनुपात या प्रतिशत के रूप में दिखाया जा सकता है।

सबसे अधिक प्रचलित विचरण का गुणांक मानक विचलन और समांतर माध्य का अनुपात है। सांकेतिक रूप में $\frac{\sigma}{m}$ विचरण का गुणांक होता है, जहाँ σ मानक विचलन है तथा m चर का समांतर माध्य है। अनुपात एक शुद्ध संख्या है। इसे प्रतिशत रूप में इस प्रकार दिखाया जा सकता है :

$$\frac{\sigma}{m} \times 100$$

यदि हम परास का प्रयोग परिक्षेपण के मापक के रूप में करें जहाँ X_{max} चर का सबसे बड़ा तथा X_{min} सबसे छोटा मूल्य है, तो हम 'विचरण के गुणांक' को निम्नलिखित विधि से भी निकाल सकते हैं।

$$\frac{X_{max}-X_{min}}{X_{max}+X_{min}} = \frac{\text{Range}}{X_{max}+X_{min}}$$

यदि हम चतुर्थक विचलन का प्रयोग परिक्षेपण के मापक के रूप में करें, तो इसे इस विधि से निकाला जा सकता है :

$$\frac{Q_3-Q_1}{Q_3+Q_1}$$

इसी प्रकार यदि हम माध्य विचलन का प्रयोग करें तो विचरण के गुणांक निम्नलिखित विधि से निकालें जा सकते हैं —

माध्य विचलन/समांतर माध्य

या,

माध्य विचलन/मध्यिका

अभ्यास

1. परिक्षेपण पद की सोदाहरण व्याख्या करें।
2. हम किसी मूल्य विशेष से परिक्षेपण का मापन क्यों करते हैं? क्या परास और चतुर्थक विचलन किसी मूल्य से परिक्षेपण के मापक हैं?
3. परिक्षेपण के एक अच्छे मापक की क्या विशेषताएँ हैं?'
4. 'निरपेक्ष' परिक्षेपण के चार विभिन्न मापक कौन-कौन से हैं? उनकी विशेषताएँ बताएँ।
5. 'विचरण का गुणांक परिक्षेपण का **सापेक्ष मापक** है।' हम विचरण के गुणांक का परिकलन परिक्षेपण के किसी भी मापक द्वारा कर सकते हैं, जैसे परास, चतुर्थक विचलन, माध्य विचलन तथा मानक विचलन। इन स्थितियों में विचरण के गुणांक के प्रयोग के उदाहरण दें।
6. निम्नलिखित मूल्यों के समांतर माध्य तथा मानक विचलन का परिकलन करें :
   क. बिना वर्गीकरण किए,
   ख. 140-145, 145-150, ...आदि वर्गों में वर्गीकृत करके
   ग. 140-150, 150-160, ...आदि वर्गों में वर्गीकृत करके

140	143	143	146	146
146	154	156	159	162
164	164	166	166	167
167	168	168	169	169
169	171	175	175	176
176	178	180	182	182
182	182	182	183	184
186	188	190	190	191
191	192	195	202	227

7. किसी परीक्षा में किसी कक्षा के 15 विद्यार्थियों ने गणित और अर्थशास्त्र में निम्नलिखित अंक प्राप्त किए :

गणित

31	25	78	65	80
56	58	42	39	54
60	63	58	45	72

अर्थशास्त्र

42	39	45	49	58
56	49	44	60	62
44	50	51	38	40

अंकों के परास, माध्य विचलन (समांतर माध्य से) तथा मानक विचलन विधियों द्वारा जाँच करें कि विद्यार्थियों के गणित के अंकों में अर्थशास्त्र के अंकों की तुलना में विचरण अधिक है?
यदि आप परिक्षेपण के किसी सापेक्ष मापक का प्रयोग करें तो क्या आपका परिणाम बदल जाएगा?

8. 50 खेतों में एक क्विंटल गेहूँ के उत्पादन की लागत (रु. में) का वितरण निम्नलिखित है :

लागत (रु. में)	**खेतों की संख्या**
40-50	3
50-60	6
60-70	12
70-80	18
80-90	9
90-100	2
	50

क. निम्नलिखित विधियों द्वारा विचरण-मापांक का परिकलन करें :
   i. प्रत्यक्ष विधि द्वारा
   ii. पद-विचलन विधि द्वारा

   तथा परिणाम की तुलना समांतर माध्य से माध्य विचलन निकाल कर करें।

ख. विचरण के गुणांक का परिकलन निम्नलिखित विधियों द्वारा करें :
   i. लागतों के मानक विचलन द्वारा ।
   ii. समांतर माध्य से लागतों के माध्य विचलन द्वारा।

   दोनों की तुलना करें। लागतों के विचरण के संबंध में आपके विचार क्या हैं?

9. चर x के निम्नलिखित मूल्य हैं :

   7, 9, 18, 11, 10, 8, 17, 13, 11, 16

   क. निम्नलिखित का परिकलन करें :
   i. x का समांतर माध्य $(\bar{x})$
   ii. x के मूल्यों का मानक विचलन
   iii. $\bar{x}$ से x के मूल्यों का माध्य विचलन

ख. निम्नलिखित का परिकलन भी करें :

i. $\Sigma(x_i - 10)^2$

ii. $\Sigma|(x_i -$ मध्यिका$)|$

ग. जाँच करें कि —

i. $\Sigma(x_i - 10)^2 > \Sigma(x_i - \bar{x})^2$

ii. $\Sigma(x_i - \bar{x}) > \Sigma|(x_i -$ मध्यिका$)|$

10. किसी सर्वेक्षण में यह देखा गया कि दूध का औसत प्रति व्यक्ति उपभोग 5 लीटर प्रति दिन है तथा विचरण का गुणांक (% के रूप में) 20 है। प्रति व्यक्ति दूध के उपभोग का **विचरण मापांक** क्या होगा?

11. वर्ष 2000 में किसी परीक्षा में 1000 विद्यार्थियों के परीक्षाफल के अनुसार औसत अंक 50% तथा मानक विचलन 3% है। वर्ष 2001 में ऐसे ही एक अध्ययन द्वारा पता चला कि औसत अंक 55% तथा मानक विचलन 5% है। क्या एक वर्ष में परीक्षाफल में सुधार हुआ है?

12. क. मान लें कि चर x, 1 से 10 के बीच कोई भी समाकल मूल्य धारण करता है। x के मूल्यों के लिए निम्नलिखित का परिकलन करें :

i. समांतर माध्य

ii. मानक विचलन

iii. समांतर माध्य से मानक विचलन

iv. मध्यिका से मानक विचलन

ख. निम्नलिखित के द्वारा x के मूल्यों के विचरण गुणांक का भी परिकलन करें :

i. मानक विचलन

ii. माध्य विचलन

13. 'व्यक्तियों के एक समूह के लिए इंचों में मापा गया ऊँचाई का मानक विचलन, फीट में मापे गए ऊँचाई के मानक विचलन से अधिक होगा।' यह कथन सही है या गलत, उचित व्याख्या द्वारा समझाएँ।

14. भारत के 55 गाँवों की जनसंख्या का आवृत्ति-वितरण नीचे दिया गया है :

जनसंख्या	**गाँवों की संख्या**
200 से कम	5
200-400	14
400-1000	21
1000-2000	9
2000-5000	3
5000 से अधिक	3
कुल	**55**

परिक्षेपण के एक उपयुक्त मापक का परिकलन करें तथा अपने चुनाव का कारण बताएँ।

अध्याय - 8

सहसंबंध का गुणांक

पिछले अध्यायों में हमने एकविचर वितरण का अध्ययन किया, जहाँ एकल चर से संबद्ध प्रेक्षण दिए गए थे। इसके द्वारा हमें चरों के मूल्यों के एक समुच्चय-विशेष के लिए केंद्रीय प्रवृत्ति के विभिन्न मापों (समांतर माध्य, मध्यिका आदि) तथा परिक्षेपण के विभिन्न मापों (मानक विचलन, माध्य विचलन आदि) की जानकारी मिली।

अब मान लें कि हमारे पास कई व्यक्तियों के बारे में **दो** चरों x तथा y से संबद्ध प्रेक्षण हैं। हमें x तथा y का एक **द्विविचर** वितरण दिया गया है।

ऐसी स्थिति में हम प्रत्येक चर x तथा y के लिए केंद्रीय मूल्य (समांतर माध्य, मध्यिका आदि) तथा **परिक्षेपण** (मानक विचलन, माध्य विचलन आदि) का परिकलन अलग-अलग कर सकते हैं।

अब हम यह भी जानना चाहते हैं कि क्या दो चरों के बीच कोई **साहचर्य** है? उदाहरण के लिए, हम जानना चाहते हैं कि दोनों में से किसी एक चर में एक निश्चित परिमाण में परिवर्तन हो (बढ़ना या घटना) तो इसके परिणामस्वरूप दूसरे चर में वृद्धि या कमी उसी अनुपात में होती है? क्या दोनों में परिवर्तन समान दिशा तथा अनुपात में होता है या पहले चर में परिवर्तन दूसरे चर में परिवर्तन के अनुपात में अधिक या कम है या एक-दूसरे की विपरीत दिशा में है?

दो चरों के बीच साहचर्य का संख्यात्मक मापन कार्ल पिअरसन (Karl Pearson) के **सहसंबंध के गुणांक** द्वारा दिया गया है। सबसे पहले हम दो चरों के बीच **साहचर्य के स्वरूप** का अध्ययन **प्रकीर्ण आरेख** के द्वारा करेंगे।

1. प्रकीर्ण आरेख (Scatter Diagram)

प्रकीर्ण आरेख में, हम दो चरों के मूल्यों को ग्राफ पत्र पर बिंदुओं के एक समुच्चय के रूप में प्लॉट करते है ....। इससे बिंदुओं का जो समुच्चय प्राप्त होता है उसे **प्रकीर्ण आरेख** कहते हैं। इसे निम्नलिखित उदाहरण द्वारा स्पष्ट किया जा सकता है।

सारणी - 8.1

पाँच ग्रामीण परिवारों के मासिक आय तथा खाद्य-सामग्रियों पर व्यय (रु. में)

चर	परिवार				
	1	**2**	**3**	**4**	**5**
आय (x)	550	600	800	700	650
खाद्य-सामग्रियों पर व्यय (y)	400	450	550	550	400

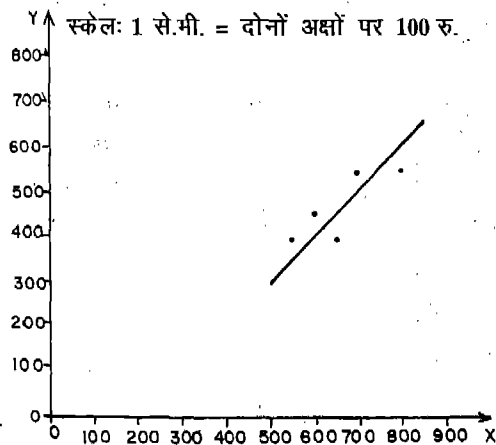
उदाहरण 1

मान लें कि हमारे पास पाँच ग्रामीण परिवारों से संबद्ध निम्नलिखित प्रेक्षण (रु. में) हैं, जो सारणी 8.1 में दिखाए गए हैं :

क. मासिक आय (x), तथा

ख. खाद्य सामग्री पर कुल मासिक व्यय (y),

हम देखते हैं कि पहले परिवार की आय 550 रु. प्रति माह है तथा खाद्य-सामग्रियों पर इस परिवार का व्यय 400 रु. प्रति माह है। हम इसे ग्राफ-पेपर पर एक बिंदु (x, y) के रूप में दिखा सकते हैं, जहाँ x = 550 तथा y = 400 है। हम 550 रु. को x-अक्ष पर तथा 400 रु. को y-अक्ष पर मापते हैं। इसी प्रकार, दूसरे परिवार के लिए x=600 है तथा y = 450। अतः दूसरे बिंदु पर कोटियाँ (600, 450) हैं। हम 600 को x-अक्ष पर तथा 450 को y-अक्ष पर मापते हैं, आदि।

बिंदुओं का एकत्रीकरण आरेख 8.1 में दिखाया गया है। इसे प्रकीर्ण-आरेख कहते हैं।

आरेख 8.1 : प्रकीर्ण आरेख (सारणी 9.1 के आँकड़ों के लिए)

प्रकीर्ण-आरेख से यह स्पष्ट होता है कि सभी बिंदु धनात्मक ढाल वाली (नीचे से ऊपर की ओर) सरल रेखा के आसपास एकत्रित हैं। दूसरे शब्दों में, जैसे-जैसे परिवार की आय (x) बढ़ती है, खाद्य-सामग्री पर व्यय (y) बढ़ता जाता है।

सामान्यतः यदि सरल-रेखा धनात्मक ढाल वाली है, तो x तथा y के मूल्य में परिवर्तन एक ही दिशा में होता है, (जैसे जब x का मूल्य बढ़ता है तो y का मूल्य भी बढ़ता है।) अन्यथा यदि सरल रेखा ऋणात्मक ढाल वाली (ऊपर से नीचे की ओर) हो तो x के मूल्य में वृद्धि के साथ y के मूल्य में कमी होगी।

सरल रेखा की **ढाल** (जो सरल रेखा के x-अक्ष पर बनाए गए कोण पर निर्भर करती है तथा जो $\frac{y}{x}$ के बराबर होती है) **x के मूल्य में एक इकाई परिवर्तन के कारण y के मूल्य में परिवर्तन की दर निर्धारित करती है।**

इसे हम निम्नलिखित उदाहरण द्वारा स्पष्ट कर सकते हैं।

उदाहरण 2

निम्नलिखित सारणी में किसी थोक-बाजार में विभिन्न मौसमों में आलू की कीमत (p) तथा माँग की मात्रा (q) दी गई है –

सारणी - 8.2

आलू की माँग

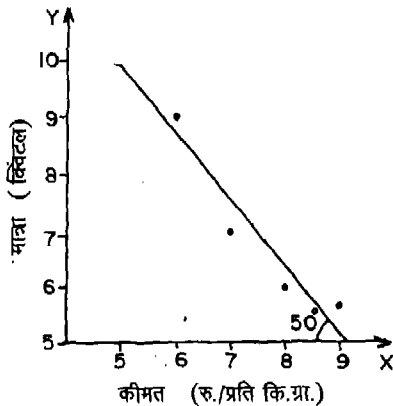
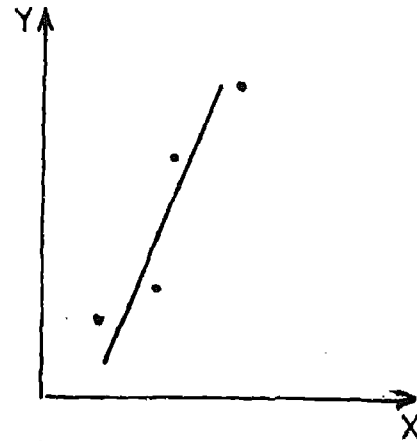
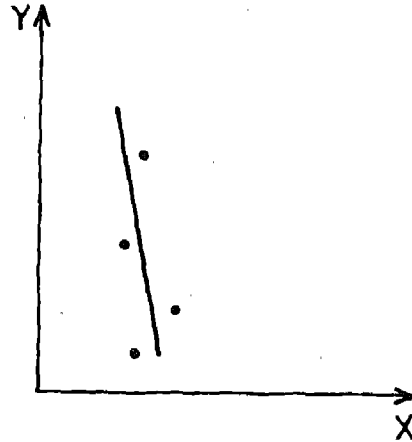
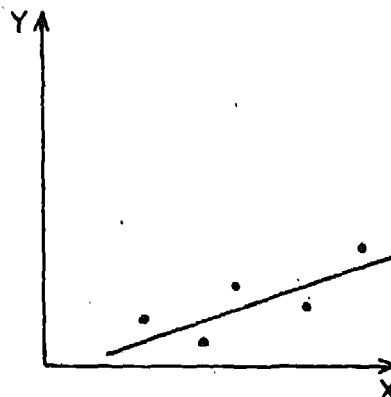
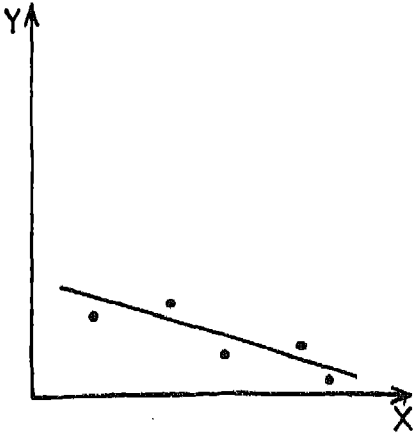
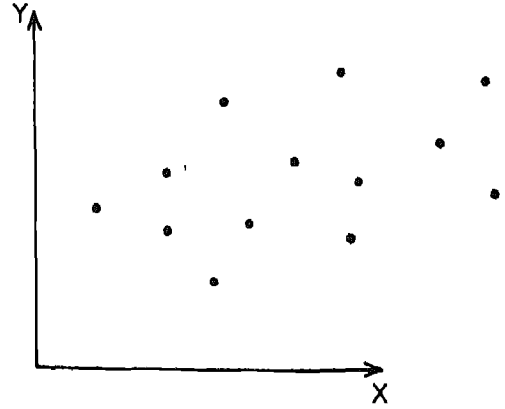
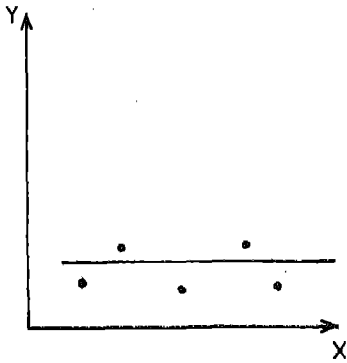
कीमत (p) (प्रति कि. ग्रा./ रु. में)	5	6	7	8	9
मात्रा (q) (क्विंटल में)	10	9	7	5	4

उपर्युक्त आँकड़ों के आधार पर एक प्रकीर्ण आरेख बनाएँ तथा p तथा q के बीच साहचर्य की प्रकृति निर्धारित करें।

प्रकीर्ण आरेख, नीचे आरेख 8.2 में दिखाया गया है। x-अक्ष पर कीमत (p) तथा y-अक्ष पर मात्रा (q) मापी गई है। q तथा p के बीच प्रतिकूल संबंध है, क्योंकि प्रकीर्ण आरेख में सभी बिंदु

ऋणात्मक ढाल वाली सरल रेखा के आसपास एकत्रित हैं। यह सरल रेखा x-अक्ष पर लगभग 50° का एक कोण बनाती है (जैसा आरेख में दिखाया गया है)। अतः माँग की गई मात्रा में कमी लगभग उसी अनुपात में होगी, जिस अनुपात में कीमत में वृद्धि होगी।

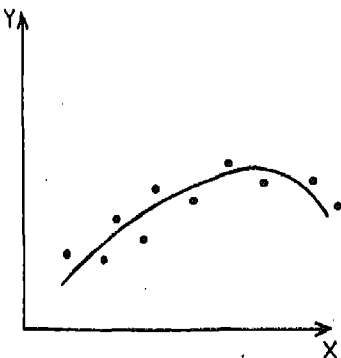
सामान्यतः, यदि सरल रेखा x-अक्ष पर 45° का कोण बनाती है, y के मूल्य में भी उसी अनुपात में परिवर्तन होता है, जिस अनुपात में x के मूल्य में।

y के मूल्य में परिवर्तन x के मूल्य में परिवर्तन के अनुपात में तब अधिक होता है जब सरल रेखा द्वारा x-अक्ष पर बनाया गया कोण 45° से अधिक होता है (तथा अनुपात में तब कम होता है जब सरल रेखा द्वारा x-अक्ष पर बनाया गया कोण 45° से कम होता है)। काल्पनिक आरेखों 8.3, 8.4, 8.5 तथा 8.6 को देखें। आरेख 8.7 यह दिखाता है कि x के मूल्य में वृद्धि के कारण y के मूल्य में कोई परिवर्तन नहीं होता (इस अंतिम स्थिति का उदाहरण है नमक जैसी किसी वस्तु की माँग की गई मात्रा का इसकी कीमत के साथ संबंध)। आरेख 8.8 x तथा y के बीच एक अरैखिक संबंध दिखाता है तथा आरेख 8.9 यह दिखाता है कि इनके बीच कोई संबंध नहीं है।

आरेख 8.2 : आनुपातिक परिवर्तन

आरेख 8.3 : अनुपात से अधिक परिवर्तन

आरेख 8.4 : अनुपात से अधिक परिवर्तन

आरेख 8.5 : अनुपात से कम परिवर्तन

आरेख 8.6 : अनुपात से कम परिवर्तन

आरेख 8.7 : कोई परिवर्तन नहीं

आरेख 8.8 : अ-रैखिक संबंध

Y

X

आरेख 8.9 : कोई संबंध नहीं

2. प्रकीर्ण आरेख के गुण व दोष

गुण

क. प्रकीर्ण आरेख को रेखांकित करना आसान है।

ख. दो चरों के बीच किस प्रकार का साहचर्य है (रैखिक या अरैखिक), इसे बताने के लिए यह एक पहला आसान कदम है।

ग. यदि x तथा y के बीच रैखिक संबंध है तो, यह x के मूल्य में परिवर्तन के कारण y के मूल्य में हुए आनुपातिक परिवर्तन का एक स्पष्ट चित्र प्रस्तुत करता है।

दोष

क. प्रकीर्ण आरेख के द्वारा साहचर्य की शक्ति (strength) या कोटि को संख्यात्मक रूप से मापना संभव नहीं।

ख. प्रकीर्ण आरेख कारण-कार्य संबंध की दिशा नहीं बताता। यह नहीं बताता कि y के कारण x में परिवर्तन होता है या x के कारण y में।

ग. यदि चरों की संख्या दो से अधिक हो तो ग्राफ-पेपर पर प्रकीर्ण आरेख बनाना संभव नहीं।

3. कार्ल पियरसन का सहसंबंध गुणांक — गुणन आघूर्ण सहसंबंध (Karl Pearson Coefficient of Correlation — Product Moment Correlation)

यदि दो चरों x तथा y के बीच एक **रैखिक** संबंध है तो हम उनके बीच के **'साहचर्य की कोटि'** को कार्ल पियरसन के सहसंबंध गुणांक द्वारा संख्यात्मक रूप से माप सकते हैं। इसे **गुणन-आघूर्ण सहसंबंध** भी कहते हैं।

मान लें कि $x_1, x_2, ...x_n$ आदि x के मान हैं तथा $y_1, y_2, ...y_n$ आदि y के संगत मान हैं। x तथा y के समांतर माध्य निम्न है :

$$\bar{x}=\frac{1}{n}\Sigma x \quad \text{तथा} \quad \bar{y}=\frac{1}{n}\Sigma y$$

तथा इनके विचलन हैं :

$$\sigma_X^2=\frac{1}{n}\Sigma(X-\bar{X})^2=\frac{1}{n}\Sigma X^2-\bar{X}^2, \text{ तथा}$$

$$\sigma_Y^2=\frac{1}{n}\Sigma(Y-\bar{Y})^2=\frac{1}{n}\Sigma Y^2-\bar{Y}^2,$$

x तथा y के मानक विचलन क्रमशः σ_x तथा σ_y उनके विचरण मापांकों के धनात्मक वर्गमूल हैं।

x तथा y के उनके माध्यों से विचलन निम्नलिखित हैं :

$$x = X - \bar{X} \text{ तथा } y \; Y \; \bar{Y}$$

x तथा y के गुणन आघूर्ण सहसंबंध को निम्न प्रकार से परिभाषित किया जाता है :

$$r=\frac{\Sigma xy}{n\sigma_x\sigma_y}$$

जो इस प्रकार भी व्यक्त किए जा सकते हैं :

$$r=\frac{\Sigma xy-\frac{1}{n}(\Sigma x)(\Sigma y)}{\sqrt{\Sigma x^2-\frac{1}{n}(\Sigma x)^2}\sqrt{\Sigma y^2-\frac{1}{n}(\Sigma y)^2}}$$

अथवा

$$r=\frac{n\Sigma xy-(\Sigma x)(\Sigma y)}{\sqrt{n\Sigma x^2-(\Sigma x^2)}\sqrt{n\Sigma y^2-(\Sigma y^2)}}$$

अब इन सूत्रों के प्रयोग द्वारा सहसंबंध गुणांक के परिकलन को स्पष्ट करें।

उदाहरण 3

उदाहरण 8.2 में दिए गए आँकड़ों के आधार पर, आलुओं की 'माँग की गई मात्रा' तथा 'कीमत' के बीच गुणन आघूर्ण सहसंबंध का परिकलन करें।

सारणी 8.3 में प्रत्यक्ष-विधि द्वारा तथा सारणी 8.4 में वास्तविक माध्य से विचलन द्वारा परिकलन दिखाए गए हैं।

सारणी - 8.3

सहसंबंध गुणांक का परिकलन

मूल्य (p)	मात्रा (q)	p^2	q^2	pq
5	10	25	100	50
6	9	36	81	54
7	7	49	49	49
8	5	64	25	40
9	4	81	16	36
$\Sigma p=35$	$\Sigma q=35$	$\Sigma p^2=255$	$\Sigma q^2=271$	$\Sigma pq=229$

हम इस सूत्र का प्रयोग करते हैं :

$$r=\frac{\Sigma pq-\frac{1}{n}(\Sigma p)(\Sigma q)}{\sqrt{\Sigma p^2-\frac{1}{n}(\Sigma p)^2}\sqrt{\Sigma q^2-\frac{1}{n}(\Sigma q)^2}}$$

इसे **प्रत्यक्ष विधि** कहते हैं। मूल्यों को प्रतिस्थापित करने पर हमें निम्नलिखित प्राप्त होता है,

$$r=\frac{229-\frac{1}{5}\times 35\times 35}{\sqrt{255-\frac{1}{5}(35)^2}\sqrt{271-\frac{1}{5}(35)^2}}=-0.99$$

उदाहरण 4

उदाहरण 2 में दिए गए आँकड़ों के आधार पर p तथा q के बीच उनके निजी माध्यों से विचलनों के द्वारा गुणन आघूर्ण सहसंबंध गुणांक का परिकलन करें।

परिकलन नीचे सारणी 8.4 में दिया जा रहा है।

अतः $\bar{p}=\frac{1}{5}\Sigma p=7$ तथा $\bar{q}=\frac{1}{5}\Sigma p=7$

सूत्र के प्रयोग द्वारा –

$$r=\frac{\Sigma(p-\bar{p})(q-\bar{q})}{\sqrt{\Sigma(p-\bar{p})^2}\sqrt{\Sigma(q-\bar{q})^2}}=\frac{-16}{\sqrt{10}\sqrt{26}}=-0.99$$

इसे **वास्तविक माध्य विधि** कहते हैं।

उदाहरण 5

पाँच ग्रामीण परिवारों के x = मासिक आय (रु. में) तथा y = मासिक व्यय (रु. में) सारणी 8.1 में दिए गए हैं।

x तथा y के बीच गुणन आघूर्ण सहसंबंध का परिकलन **प्रत्यक्ष विधि** द्वारा करें।

सारणी - 8.4

सहसंबंध का गुणांक

(वास्तविक माध्यों से विचलन)

p	q	$p-\bar{p}$	$q-\bar{q}$	$(p-\bar{p})^2$	$(q-\bar{q})^2$	$(p-\bar{p})(q-\bar{q})$
5	10	-2	3	4	9	-6
6	9	-1	2	1	4	-2
7	7	0	0	0	0	0
8	5	1	-2	1	4	-2
9	4	2	-3	4	9	-6
$\Sigma p=35$	**$\Sigma q=35$**			**$\Sigma(p-\bar{p})^2=10$**	**$\Sigma(q-\bar{q})^2=26$**	**$\Sigma(p-\bar{p})(q-\bar{q})=16$**

सारणी - 8.5

सहसंबंध का परिकलन

X	Y	X^2	Y^2	XY
550	400	302500	160000	220000
600	450	360000	202500	270000
800	550	640000	302500	440000
700	550	490000	302500	385000
650	400	422500	160000	260000
$\Sigma X=3300$	**$\Sigma Y=2350$**	**$\Sigma X^2=2215000$**	**$\Sigma Y^2=1127500$**	**$\Sigma XY=1575000$**

परिकलन सारणी 8.5 में दिया गया है :

सूत्र का प्रयोग करते हुए :

$$r=\frac{\Sigma XY-\frac{1}{n}(\Sigma X)(\Sigma Y)}{\sqrt{\Sigma X^2-\frac{1}{n}(\Sigma X)^2}\sqrt{\Sigma Y^2-\frac{1}{n}(\Sigma Y)^2}}$$

$$=\frac{1575000-\frac{1}{5}(3300)(2350)}{\sqrt{2215000-\frac{1}{5}(3300)^2}\sqrt{1127500-\frac{1}{5}(2350)^2}}$$

$$=\frac{24000}{\sqrt{37000\times 23000}}=\frac{24000}{29171.9}$$

$$=0.82$$

उदाहरण 6

पाँच परिवारों के x (मासिक-आय रु. में) तथा y (खाद्य-सामग्रियों पर मासिक व्यय, रुपयों में) सारणी 8.1 में दिए गए हैं।

x तथा y के बीच गुणन आघूर्ण सहसंबंध **वास्तविक माध्यों से उनके विचलनों** द्वारा निकालें।

परिकलन नीचे सारणी 8.6 में दिए गए हैं।

$$\bar{x}=\frac{1}{n}\Sigma x=660, \bar{y}=\frac{1}{n}\Sigma y=470$$

सूत्र का प्रयोग करते हुए,

$$r=\frac{\Sigma xy}{\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}}=\frac{24000}{\sqrt{37000\times 23000}}$$

$$=\frac{24000}{29171.9}=0.82$$

4. सहसंबंध गुणांक के परिकलन के लिए पद-विचलन विधि का प्रयोग

सहसंबंध गुणांक का परिकलन काफी आसान हो जाएगा यदि —

क. हम x तथा y का विचलन **काल्पनिक मानों** से करें, (इसे **'उद्गम-परिवर्तन'** कहते है), तथा

ख. विचलनों को किसी सुविधाजनक मान या समापवर्तक से भाग दें (इसे **'स्केल परिवर्तन'** कहते हैं)।

अब हम इसे ऊपर दिए गए उदाहरणों से स्पष्ट करेंगे। नीचे एक महत्त्वपूर्ण निष्कर्ष दिया गया है :

सहसंबंध गुणांक का मूल्य x तथा y के उद्गम परिवर्तन तथा स्केल-परिवर्तन से प्रभावित नहीं होता।

सारणी - 8.6

सहसंबंध का परिकलन
(वास्तविक माध्यों से विचलन)

X	Y	$X-\bar{X}=x$	$Y-\bar{Y}=y$	x^2	y^2	xy
550	400	-110	-70	12100	4900	7700
600	450	-60	-20	3600	400	1200
800	550	140	80	19600	6400	11200
700	550	40	80	1600	6400	3200
650	400	-10	-70	100	4900	700
Σx = 3300	**Σy = 2350**			**37000**	**23000**	**24000**

उदाहरण 7

सारणी 8.1 में पाँच ग्रामीण परिवारों से संबद्ध निम्नलिखित आँकड़े दिए गए हैं :

x = मासिक आय (रुपयों में), तथा

y = मासिक व्यय (रुपयों में)

x तथा y के बीच गुणन आघूर्ण सहसंबंध का परिकलन करें।

पद-विचलन विधि द्वारा यह परिकलन सारणी 8.7 में दिखाया गया है। इसमें हम 600 को x का तथा 550 को y का काल्पनिक माध्य मानते हैं तथा इन दोनों माध्यों से किए गए विचलनों को समापवर्तक 50 से भाग देते हैं।

$$d'_x = \frac{x-600}{50} \text{ and } d'_y = \frac{y-550}{50}$$

सूत्र का प्रयोग करते हुए,

$$r = \frac{N\sum d'_x d'_y - \sum d'_x \sum d'_y}{\sqrt{N\sum d'^2_x - (\sum d'_x)^2}\sqrt{N\sum d'^2_y - (\sum d'_y)^2}}$$

$$= \frac{5\times0-6\times(-8)}{\sqrt{5\times22-6^2}\sqrt{5\times22-(-8)^2}} = 0.82$$

उदाहरण 8

नीचे दी गई सारणी 8.8 के आँकड़ों के लिए निम्नलिखित के बीच सहसंबंध गुणांक का परिकलन करें, जहाँ —

x = स्कूल जाने के वर्ष

y = प्रति एकड़ वार्षिक उत्पादन (रु. में)

सारणी - 8.8

X	0	2	4	6	8	10	12	14	14	16
Y	4	4	6	10	10	8	12	10	8	6

तुलना करने के लिए हम सहसंबंध गुणांक का परिकलन निम्नलिखित तीन विधियों द्वारा कर सकते हैं :

क. प्रत्यक्ष विधि द्वारा (सारणी 8.9)

ख. काल्पनिक माध्य से विचलन द्वारा (सारणी 8.10)

ग. पद-विचलन विधि द्वारा (सारणी 8.11)।

सारणी - 8.9

सहसंबंध गुणांक का परिकलन

(प्रत्यक्ष विधि)

X	Y	X^2	Y^2	XY
0	4	0	16	0
2	4	4	16	8
4	6	16	36	24
6	10	36	100	60
8	10	64	100	80
10	8	100	64	80
12	12	144	144	Z144
14	10	196	100	140
14	8	196	64	112
16	6	256	36	96
$\Sigma X=86$	$\Sigma Y=78$	$\Sigma X^2=1012$	$\Sigma Y^2=676$	$\Sigma XY=744$

सारणी - 8.7

सहसंबंध गुणांक का परिकलन

(पद विचलन विधि)

X	Y	$d'_x = \frac{X-600}{50}$	$d'_y = \frac{Y-550}{50}$	d'^2_x	d'^2_y	$d'_x d'_y$
550	400	-1	-3	1	9	3
600	450	0	-2	0	4	0
800	550	4	0	16	0	0
700	550	2	0	4	0	0
650	400	1	-3	1	9	-3
		$\Sigma d'_x = 6$	$\Sigma d'_y = 8$	$\Sigma d'^2_x = 22$	$\Sigma d'^2_y = 22$	$\Sigma d'_x d'_y = 0$

सूत्र का प्रयोग करते हुए,

$$r=\frac{\sum XY-n\bar{X}\bar{Y}}{\sqrt{\sum X^2-n\bar{X}^2}\sqrt{\sum Y^2-n\bar{Y}^2}}$$

$$=\frac{744-10\times8.6\times7.8}{\sqrt{1012-10\times8.6^2}\sqrt{676-10\times7.8^2}}$$

$$=\frac{73.2}{\sqrt{272.4}\sqrt{67.6}}=\frac{73.2}{135.7}=0.54$$

सारणी - 8.10

सहसंबंध का परिकलन

(वास्तविक माध्यों से विचलन)

$x=X-\bar{X}$	$y=Y-\bar{Y}$	x^2	y^2	xy
-8.6	-3.8	73.96	14.44	32.68
-6.6	-3.8	43.56	14.44	25.08
-4.6	-1.8	21.16	3.24	8.28
-2.6	2.2	6.76	4.84	-5.72
-0.6	2.2	0.36	4.84	-1.32
1.4	0.2	1.96	.04	0.28
3.4	4.2	11.56	17.64	14.28
5.4	2.2	29.16	4.84	11.88
5.4	0.2	29.16	.04	1.08
7.4	-1.8	54.76	3.24	-13.32
$\sum x=0$	$\sum y=0$	$\sum x^2=272.4$	$\sum y^2=67.60$	$\sum xy=73.20$

सूत्र का प्रयोग करते हुए,

$$r=\frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}=\frac{73.20}{\sqrt{272.4\times67.6}}=0.54$$

सारणी - 8.11

सहसंबंध गुणांक का परिकलन

(पद-विचलन विधि)

(x का वास्तविक माध्य 10 तथा y का वास्तविक माध्य 8 है; समापवर्तक 2 है)

X	$d'_x=\frac{X-10}{2}$	d'^2_x	Y	$d'_y=\frac{Y-8}{2}$	d'^2_y	$d'_x d'_y$
0	-5	25	4	-2	4	10
2	-4	16	4	-2	4	8
4	-3	9	6	-1	1	3
6	-2	4	10	1	1	-2
8	-1	1	10	1	1	-1
10	0	0	8	0	0	0
12	1	1	12	2	4	2
14	2	4	10	1	1	2
14	2	4	8	0	0	0
16	3	9	6	-1	1	-3
	$\sum d'_x=7$	$\sum d'^2_x=73$		$\sum d'_y=1$	$\sum d'^2_y=17$	$\sum d'_x d'_y=19$

$$r_{xy}=r_{d'_x d'_y}=\frac{\sum d'_x d'_y}{\sqrt{\sum d'^2_x}\sqrt{\sum d'^2_y}}=\frac{19}{\sqrt{73\times17}}=0.54$$

हम देखते हैं कि तीनों विधियों के द्वारा समान परिणाम प्राप्त होते हैं।

5. सहसंबंध गुणांक की विशेषताएँ

सहसंबंध गुणांक की निम्नलिखित विशेषताएँ ध्यान देने योग्य हैं :

क. **सहसंबंध गुणांक कभी –1 से कम तथा +1 से अधिक नहीं होता $-1 \le r \le +1$.**

ख. **r का ऋणात्मक मूल्य दो चरों X तथा Y के बीच एक विलोम संबंध दर्शाता है।** X के मूल्य में वृद्धि के साथ Y के मूल्य में कमी होती है तथा X के मूल्य में कमी के साथ Y के मूल्य में वृद्धि होती है। उदाहरण के लिए, किसी वस्तु की माँग की गई मात्रा तथा उसकी कीमत में विपरीत संबंध है।

ग. **यदि X के मूल्य में वृद्धि के साथ Y के मूल्य में भी वृद्धि होती है तो r धनात्मक होगा।** उदाहरण के लिए, आय तथा माँग की गई मात्रा में धनात्मक संबंध होता है।

घ. **यदि r=0 हो, तो दो चरों X तथा Y के बीच कोई संबंध नहीं है।** इस स्थिति में एक चर के मूल्य में परिवर्तन का दूसरे चर के मूल्य में परिवर्तन से रैखिक संबंध नहीं है।

ङ. **यदि r = + 1, या r = -1 हो, तो दो चरों X तथा Y के बीच पूर्ण रैखिक संबंध है।** X तथा Y के बीच संबंध यथातथ है।

च. **r का ऊँचा मूल्य (+1 या -1 के निकट) X तथा Y के बीच निकट संबंध दर्शाता है।**

छ. **r का मूल्य यदि शून्य (0) के निकट हो (धनात्मक या ऋणात्मक) तो यह X तथा Y के बीच कमजोर रैखिक संबंध दिखाता है।**

ज. **r का मूल्य 'उद्‌गम परिवर्तन' या 'स्केल परिवर्तन' से प्रभावित नहीं होता।**

$$d'_x = \frac{X-A}{B} \text{ तथा } d'_y = \frac{Y-C}{D}$$

जहाँ A तथा C काल्पनिक माध्य हैं तथा B व D कोई मूल्य हैं (या समापवर्तक हैं) तो

$$rd'_x d'_y = r_{x,y}$$

सहसंबंध गुणांक का एक दोष है कि यह कारण की दिशा नहीं बताता, जैसे यह नहीं बताता कि X में परिवर्तन का कारण Y है या Y में परिवर्तन का कारण X है।

6. स्पियरमैन का कोटि सहसंबंध (Spearman's Rank Correlation)

कभी-कभी चरों को संख्यात्मक रूप से मापना संभव नहीं होता। उदाहरण के लिए, लोगों की बुद्धि या उनके शारीरिक सौन्दर्य या कला व संगीत में उनकी रुचि आदि को उस प्रकार संख्यात्मक रूप से नहीं मापा जा सकता जैसे उनके वजन या कद को। ऐसे चरों को **गुण** कहते हैं। हम व्यक्तियों को इन गुणों के अनुसार कोटियों में बाँट सकते हैं। इन कोटियों का प्रयोग विश्लेषण के उद्देश्य से किया जाता है।

मान लें कि n व्यक्तियों को उनकी बुद्धि तथा शारीरिक सौन्दर्य के अनुसार कोटियों में बाँटा गया जो नीचे दी गई हैं :

	व्यक्ति			
	पहला	**दूसरा**	.....	**n वाँ**
बुद्धि (R)	R_1	R_2	.....	R_n
शारीरिक सौन्दर्य (R')	R_1'	R_2'	.....	R_n'

स्पियरमैन का कोटि-सहसंबंध-गुणांक व्यक्तियों अथवा वस्तुओं की कोटियों के बीच रैखिक साहचर्य को उनके गुणों के प्रकारों (योग्यता) के अनुसार मापता है।

कोटि सहसंबंध के परिकलन में निम्नलिखित चरण हैं :

चरण 1 : कोटियों के बीच अंतर निकालें जहाँ प्रत्येक व्यक्ति के लिए D=R–R'

चरण 2 : सभी व्यक्तियों की कोटियों में अंतर के वर्गों का योग करके $\sum D^2$ निकालें।

चरण 3 : स्पियरमैन के कोटि-सहसंबंध को निम्न-लिखित सूत्र द्वारा प्राप्त करें?

$$r_k = 1 - \frac{6\sum D^2}{n^3 - n}$$

जहाँ n व्यक्तियों की संख्या है।

हम कोटि सहसंबंध के परिकलन के लिए निम्नलिखित स्थितियों पर विचार करेंगे :

स्थिति 1 : जब वास्तविक कोटियाँ दी गई हैं।

स्थिति 2 : जब कोटियाँ नहीं दी गई हैं, परंतु इन्हें वास्तविक मूल्यों द्वारा निकाला जा सकता है।

स्थिति 3 : जब कोटियों की पुनरावृत्ति होती है, अर्थात् जब दो या तीन कोटियाँ बराबर हो जाती हैं।

इन्हें हम उदाहरणों द्वारा स्पष्ट करेंगे।

स्थिति 1 : जब वास्तविक कोटियाँ दी गई हैं।

उदाहरण 9

पाँच विद्यार्थियों को अर्थशास्त्र तथा गणित में उनकी योग्यता के अनुसार कोटियों में बाँटा गया है। उनकी **कोटियाँ** नीचे सारणी में दी जा रही हैं :

	विद्यार्थी				
	क	**ख**	**ग**	**घ**	**ङ**
गणित में कोटियाँ	1	2	3	4	5
अर्थशास्त्र में कोटियाँ	4	2	1	3	5

स्पियरमैन के कोटि-सहसंबंध का परिकलन करें। परिकलन नीचे सारणी 8.12 में दिखाया गया है।

सारणी - 8.12

कोटि सहसंबंध का परिकलन

विद्यार्थी	गणित में कोटि (R)	अर्थशास्त्र में कोटि (R)	R–R'=D	D^2
क	1	4	-3	9
ख	2	2	0	0
ग	3	1	2	4
घ	4	3	1	1
ङ.	5	5	0	0
			$\sum D^2$=14	

कोटि सहसंबंध $r_k = 1 - \frac{6\sum D^2}{n^3-n}$,

n=5 तथा $\sum D^2$=14 प्रतिस्थापित करने के बाद हमें निम्नलिखित प्राप्त होता है :

$$r_k = 1 - \frac{6\times 14}{5^3-5} = 1 - \frac{84}{120} = 0.3$$

अतः कोटियों के बीच कमजोर लेकिन धनात्मक सहसंबंध है।

उदाहरण 10

सात विद्यार्थियों को उनकी 'मौखिक' व 'लिखित' परीक्षाओं में उनके निष्पादन के आधार पर निम्नलिखित कोटियों में बाँटा गया है :

लिखित परीक्षा में कोटियाँ	मौखिक परीक्षा में कोटियाँ
1	2
2	1
3	5
4	3
5	6
6	7
7	4

यदि ये दोनों परीक्षाएँ दो अलग-अलग परीक्षा-समितियों द्वारा ली गई हों, तो बताएँ कि आपके विचार से ये दोनों कोटियों के समुच्चय एक-दूसरे के निकट हैं या नहीं?

स्पियरमैन के कोटि-सहसंबंध के परिकलन के चरण नीचे सारणी 8.13 में दिए गए हैं।

सारणी - 8.13

कोटि सहसंबंध का परिकलन

लिखित परीक्षा में कोटि (R)	मौखिक परीक्षा में कोटि (R')	R–R'=D	D^2
1	2	-1	1
2	1	1	1
3	5	-2	4
4	3	1	1
5	6	-1	1
6	7	-1	1
7	4	3	9
			$\sum D^2$=18

अतः सूत्र $r_k = 1 - \frac{6\sum D^2}{n^3-n}$ में n = 7 तथा $\sum D^2$=18 प्रतिस्थापित करते हुए, जिससे हमें

$$r_k = 1 - \frac{6\times 18}{7^3-7} = 0.68$$ प्राप्त होता है।

चूँकि $r_k = 0.68$ सामान्यतः उच्च-सहसंबंध दिखाता है, जिससे यह निष्कर्ष निकलता है कि विद्यार्थियों की 'लिखित' एवं 'मौखिक' परीक्षाओं में प्राप्त कोटियाँ एक-दूसरे के काफी निकट हैं।

स्थिति 2: जब कोटियाँ वास्तविक मूल्यों द्वारा प्राप्त की जा सकती हैं।

इसे हम निम्नलिखित उदाहरण द्वारा स्पष्ट कर सकते हैं।

उदाहरण 11

छः विद्यार्थियों द्वारा गणित तथा अर्थशास्त्र में प्राप्त प्रतिशत अंक सारणी 8.14 में दिए गए हैं :

सारणी - 8.14

गणित तथा अर्थशास्त्र में प्रतिशत अंक

विद्यार्थी	गणित X	अर्थशास्त्र Y
क	85	60
ख	60	48
ग	55	49
घ	65	50
ड	75	55
च	90	62

X तथा Y के बीच कोटि सहसंबंध का गुणांक ज्ञात करें।

परिकलन के विभिन्न चरण सारणी 8.15 में दिखाए गए हैं। सर्वाधिक मान को पहली कोटि में रखा गया है; सर्वाधिक मान से नीचे वाले मान को दूसरी कोटि में इत्यादि। सबसे कम मूल्य को छठी कोटि में रखा गया है। यहाँ कोटियों में कोई एक-दूसरे के बराबर नही हैं।

अतः $r_k = 1 - \frac{6\sum D^2}{n^3-n} = 1-\frac{6\times12}{210} = 0.94$

यहाँ दोनों कोटियों में उच्च-सहसंबंध है।

उदाहरण 12

दिए गए आँकड़ों के आधार पर X तथा Y के बीच कोटि सहसंबंध निकालें।

X	64	63	39	40	97	31	07	84	46	82
Y	26	44	04	48	65	44	40	51	11	58

परिकलन सारणी 8.16 में दिखाया गया है।

सारणी - 8.15

कोटि सहसंबंध का परिकलन

गणित में अंक X	गणित में कोटि R	अर्थशास्त्र में अंक Y	अर्थशास्त्र में कोटि (R')	R–R'=D	D^2
85	2	60	2	0	0
60	5	48	6	-1	1
55	6	49	5	1	1
65	4	50	4	0	0
75	3	55	3	0	0
90	1	62	1	0	0
कुल					**2**

सारणी - 8.16

कोटि सहसंबंध का परिकलन

X	X के अनुसार कोटि (R)	Y	Y के अनुसार कोटि (R')	D=R–R'	D^2
64	4	26	8	–4	16
63	5	44	5	0	0
39	8	4	10	-2	4
40	7	48	4	3	9
97	1	65	1	0	0
31	9	43	6	3	9
07	10	40	7	3	9
84	2	51	3	-1	1
46	6	11	9	-3	9
82	3	58	2	1	1
कुल					**58**

पहले की तरह हम सबसे अधिक मूल्य से आरंभ करते हैं तथा इसे पहली कोटि में रखते हैं, इसके बाद वाले को दूसरी कोटि में आदि। सबसे कम मूल्य को दसवीं कोटि में रखते हैं। कोई दो कोटियाँ एक-दूसरे के बराबर नहीं हैं। अतः

$$r_k = 1 - \frac{6\sum D^2}{n^3 - n}$$

$$= 1 - \frac{6 \times 58}{10^3 - 10}$$

$$= 1 - \frac{348}{990} = 0.65$$

यहाँ ध्यान दिया जाना चाहिए कि कोटि-सहसंबंध r_k कोटियों के बीच गुणन आघूर्ण सहसंबंध के बराबर होता है।

उदाहरण के लिए, सारणी 8.16 में दी गई कोटियों के बीच गुणन आघूर्ण सहसंबंध निकालें।

परिकलन सारणी 8.17 में दिया गया है।

सारणी - 8.17

सारणी 8.16 की कोटियों के बीच गुणन आघूर्ण सहसंबंध

R	R'	R^2	R'^2	RR'
4	8	16	64	32
5	5	25	25	25
8	10	64	100	80
7	4	49	16	28
1	1	1	1	1
9	6	81	36	54
10	7	100	49	70
2	3	4	9	6
6	9	36	81	54
3	2	9	4	6
$\sum R=55$	$\sum R'=55$	$\sum R^2=385$	$\sum R'^2=385$	$\sum RR'=356$

$$r_{R,R'} = \frac{n\sum RR' - (\sum R)(\sum R')}{\sqrt{n\sum R^2 - (\sum R)^2}\sqrt{n\sum R'^2 - (\sum R')^2}}$$

$$= \frac{10 \times 356 - (55)(55)}{\sqrt{10 \times 385 - 55^2}\sqrt{10 \times 385 - 55^2}}$$

$$= 0.65$$

जो पहले की ही तरह है।

उदाहरण 9 में हमने परिकलन किया था कि

$$r_k = 1 - \frac{6\sum D^2}{n^3 - n} = 0.3$$

कोटियों के बीच गुणन आघूर्ण सहसंबंध निकालने पर हमें $r_{R,R'} = 30$ प्राप्त होता है।

स्थिति 3 : जब कोटियों की पुनरावृत्ति होती है

किसी परीक्षा में दो या अधिक विद्यार्थियों के अंक बराबर हो सकते हैं तथा उनकी कोटियों में समानता हो सकती है। इसी प्रकार विभिन्न जिलों में साक्षरता-स्तर में तुलना के क्रम में हमें ऐसे जिले भी मिल सकते हैं जहाँ साक्षरता-दर एक-दूसरे के बराबर हो। ऐसी स्थिति में इन जिलों को एक ही कोटि में रखा जाता है, आदि ।

हम निम्नलिखित उदाहरण द्वारा ऐसी स्थिति में कोटि सहसंबंध को स्पष्ट करेंगे।

उदाहरण : x तथा y के मूल्य नीचे दिए गए हैं :

X	25	45	35	40	15	19	35	42
Y	55	60	30	35	40	42	36	48

X तथा Y के बीच कोटि सहसंबंध का परिकलन करें।

X तथा Y के मूल्यों को अवरोही क्रम में व्यवस्थित करें। सर्वोच्च मूल्य को कोटि 1, दूसरे को कोटि 2 आदि में रखें। निम्नतम मूल्य को

निम्नतम कोटि 8 में रखा जाता है। हम देखते हैं कि दो स्थानों पर (चौथे और पाँचवें स्थान पर) X का मूल्य 35 है। इन दोनों को एक औसत कोटि $\left(\frac{4+5}{2}\right)$वीं $= 4.5$ वीं कोटि में रखा जा सकता है।

इसके बाद कोटियाँ सारणी 8.18 में दिखाई गई हैं। r_k का परिकलन पहले की ही तरह किया जाता है

$$r_k = 1 - \frac{6\sum D^2}{n^3-n}$$

सारणी - 8.18

श्रेणी-सहसंबंध का परिकलन

X	Y	R	R'	R–R'=D	D^2
25	55	6	2	4	16
45	60	1	1	0	0
35	30	4.5	8	-3.5	12.25
40	35	3	7	-4	16
15	40	8	5	3	9
19	42	7	4	3	9
35	36	4.5	6	-1.5	2.25
42	48	2	3	-1	1
कुल					**$\sum D^2$=65.5**

$$r_k = 1 - \frac{6\times 65.6}{8^3-8} = 1 - \frac{393}{512-8} = 1 - \frac{393}{504} = 0.22$$

अभ्यास

1. प्रकीर्ण आरेख क्या है? दो चरों x तथा y के बीच संबंध बताने में यह किस प्रकार सहायक होता है?
2. x तथा y के बीच किस प्रकार का संबंध है जब प्रकीर्ण आरेख पर सभी बिंदु —
   क. x-अक्ष के सामांतर एक सीधी रेखा के आसपास एकत्रित हों।
   ख. y-अक्ष के समांतर एक सीधी रेखा के आसपास एकत्रित हों।
   ग. ऊपर की ओर जाने वाली सरल रेखा के आसपास एकत्रित हों।
   घ. नीचे की ओर जाने वाली सरल रेखा के आसपास एकत्रित हों।
3. यदि प्रकीर्ण आरेख के सभी बिंदु ऐसी सरल रेखा के निकट एकत्र हों जो X-अक्ष पर 30° का कोण बनाती है, तो आप X तथा Y के बीच साहचर्य के बारे में क्या कहेंगे?
4. क. कार्ल पियरसन के सहसंबंध गुणांक को कैसे परिभाषित किया जाता है।
   ख. सहसंबंध गुणांक r की सीमाएँ क्या हैं?
   ग. यदि $r = +1$ या $r = -1$ हो, तो X तथा Y के बीच किस प्रकार का संबंध है?
5. क. स्पियरमैन के कोटि सहसंबंध (r_k) की परिभाषा दें।
   ख. (r_k) की सीमाएँ क्या हैं?
   ग. यदि X तथा Y के मूल्यों को कोटिबद्ध कर हम X तथा Y के बीच गुणन आघूर्ण-सहसंबंध की परिकल्पना करते हैं तो क्या यह सहसंबंध r_k के मूल्य के बराबर होगा?
6. X तथा Y के बीच सहविचरण को $\text{cov}(X,Y)= \frac{1}{n}\sum(X-\bar{X})(Y-\bar{Y})$ उसी प्रकार परिभाषित करते हैं जैसे X के विचरण तथा Y के विचरण को करते हैं। यदि विचरण (X,Y) के सूत्र के स्थान पर निम्नलिखित सहसंबंध गुणांक सूत्र का प्रयोग किया जाए तो इसके क्या लाभ हैं :

$$r = \frac{1}{n}\ \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sigma_x\sigma_y}$$

7. किसी देश के किसी राज्य में महामारी फैलने के कारण कई गाँवों में मौतें हो गई। सरकार ने इसके लिए कई कदम उठाए तथा प्रभावित गाँवों में डाक्टरों की एक टीम भेजी। बाद में किसी सांख्यविद् ने विभिन्न गाँवों में हुई मौतों की संख्या (X) तथा विभिन्न गाँवों में भेजें गए डाक्टरों की संख्या (Y) के संबंध में आँकड़े एकत्र किए तथा पाया कि $r_{x,y} = 0.8$।

   क्या आपके विचार में X तथा Y के बीच मूर्खतापूर्ण सहसंबंध है? अपने उत्तर के लिए कारण बताएँ।

8. X तथा Y के लिए निम्नलिखित आँखड़े दिए गए हैं :

X	-3	-2	-1	0	1	2	3
$Y = X^2 =$	9	4	1	0	1	4	9

   X तथा Y के बीच सहसंबंध निकालें तथा परिणामों का विश्लेषण करें।

9. दस विद्यार्थियों द्वारा 'लिखित' व 'मौखिक' परीक्षा में प्राप्त किए गए अंक नीचे दिए गए हैं :

लिखित	81	62	74	78	93	69	72	83	90	84
मौखिक	76	71	69	76	87	62	80	75	92	99

   दिए गए आँकड़ों के आधार पर —

   क. एक प्रकीर्ण आरेख बनाएँ।

   ख. यह बताएँ कि प्रकीर्ण आरेख के द्वारा अंकों के दो समुच्चयों के बीच किस प्रकार का साहचर्य प्राप्त होता है।

   ग. X तथा Y के बीच साहचर्य की शक्ति (Strength of Association) ज्ञात करने के लिए क्या आप सहसंबंध गुणांक का परिकलन करना चाहेंगे?

10. X तथा Y के मूल्य नीचे दिए गए हैं :

X	0	1	2	3	4	5
Y	10	10	10	10	10	10

   क. एक प्रकीर्ण आरेख बनाएँ तथा साहचर्य का स्वरूप निर्धारित करें।

   ख. r का मूल्य निकालें।

11. X तथा Y के लिए निम्नलिखित आँकड़े दिए गए है :

X	64	63	39	40	97	31	07	84	48	82
Y	26	44	04	48	65	44	40	51	11	58

   क. एक प्रकीर्ण आरेख बनाएँ तथा X व Y के बीच सहसंबंध के स्वरूप का निर्धारण करें।

   ख. X तथा Y के बीच सहसंबंध-गुणांक का परिकलन करें।

   ग. X तथा Y के मूल्य को कोटियों में वर्गीकृत करें तथा X व Y के बीच कोटि-सहसंबंध निकालें।

   घ. X तथा Y की (ग) में प्राप्त कोटियों के बीच गुणन आघूर्ण सहसंबंध का परिकलन करें।

12. निम्नलिखित सारणी में वर्तमान कीमतों पर भारत के निर्यात व आयात के निरपेक्ष मूल्य (हज़ार करोड़ रुपये के निकटतम) दिए गए हैं :

वर्ष	निर्यात	आयात
1996-97	119	139
1997-98	130	154
1998-99	142	176
1999-00	119	149

निर्यात तथा आयात के बीच सहसंबंध निकालें।

13. केंद्र सरकार की राजस्व प्राप्तियाँ (X) तथा कुल व्यय (Y) नीचे सारणी में दिए गए हैं –

(हज़ार करोड़ रु. में)

वर्ष	राजस्व प्राप्तियाँ*	कुल व्यय**
1990-91	55	98
1995-96	110	168
1996-97	126	190
1997-98	134	216
1998-99	151	255
1999-2000	183	284

* राजस्व प्राप्तियाँ कर व गैर-कर स्रोतों के आय के बराबर है।

** कुल व्यय में योजना तथा योजनेतर व्यय संमिलित है।

X तथा Y के बीच सहसंबंध ज्ञात करें।

14. यदि $n = 8$, $\Sigma X = 360$, $\Sigma X^2 = 20400$, $\Sigma XY = 13440$, $\Sigma Y = 272$, $\Sigma Y^2 = 9920$ हो तो दिखाएँ कि X तथा Y के बीच आघूर्ण सहसंबंध 0.7 है।

15. क. X तथा Y के बीच सह-विचरण मापांक (covariance) $cov(X,Y) = \frac{1}{n}\Sigma(X-\bar{X})(Y-\bar{Y})$ यदि ऋणात्मक हो, तो X तथा Y के बीच सहसंबंध घनात्मक होगा या ऋणात्मक?

 ख. Cov (X,Y) के मापन की इकाइयाँ क्या हैं?

 ग. X तथा Y के बीच सह-विचरण मापांक के लिए क्या +1 से अधिक या -1 से कम मूल्य प्राप्त हो सकता है?

16. निम्नलिखित सारणी दस वर्षों की अवधि के लिए निजी उपभोग व्यय तथा सकल राष्ट्रीय उत्पाद (GNP) दिखाती है :

निजी उपभोग व्यय (करोड़ रु. में)	सकल राष्ट्रीय उत्पाद (करोड़ रु. में)
11761	14950
12427	15883
12929	16992
14567	19542
17355	22896
17826	23947
20349	27380
24561	32138
24945	33291
27158	36942

क. ऊपर दिए गए आँकड़ों के आधार पर एक प्रकीर्ण आरेख बनाएँ, जिसमें सकल राष्ट्रीय उत्पाद X-अक्ष पर तथा निजी उपभोग व्यय Y-अक्ष पर दिए गए हों। प्रकीर्ण आरेख इन दो चरों के बीच किस प्रकार का संबंध दर्शाता है?

ख. दो चरों के बीच गुणन आघूर्ण सहसंबंध निकालें। क्या r का यह मूल्य ऊपर दिए गए संबंध की प्रकृति की पुष्टि करता है?

17. कुछ वर्षों के लिए मूल्य सूचकांक (P) तथा मुद्रा पूर्ति (M) नीचे दिए गए हैं :

मूल्य सूचकांक (P)	**मुद्रा पूर्ति (M) (करोड़ रु. में)**
94.5	1862
78.3	1049
100.0	2725
135.5	4237
179.7	6729

P तथा M के बीच गुणन आघूर्ण सहसंबंध निकालें। यह क्या दर्शाता है?

सूचकांकों का परिचय

1. अर्थ

सूचकांक निम्नलिखित चरों के मूल्यों में परिवर्तन की माप करते हैं —

क. वस्तुओं की कीमतें
ख. औद्योगिक उत्पादन
ग. कृषि उत्पादन
घ. निर्वाह व्यय

एक सरल उदाहरण लें। अक्तूबर 2001 में प्याज की कीमत 2 रु. 50 पैसे प्रति किलोग्राम थी जो बढ़कर दिसंबर 2001 में 3 रु. प्रति किलोग्राम हो गई। दो महीनों में प्याज की कीमत में परिवर्तन को दो विधियों द्वारा मापा जा सकता है —

i. वास्तविक अंतर

दो महीनों में प्याज की कीमत में वास्तविक अंतर 3 रु. - 2 रु.50 पैसे = 50 पैसे के बराबर है। दिसम्बर में प्याज की कीमत अक्तूबर की तुलना में 50 पैसे प्रति किलोग्राम अधिक थी।

ii. सापेक्ष अंतर

कीमतों में सापेक्ष अंतर को इस प्रकार दिखाया जा सकता है —

$$\frac{3.00-2.50}{2.50}=0.20$$

जो **मूल कीमतों की तुलना में (सापेक्ष) वर्तमान कीमतों में वास्तविक अंतर दिखाता है**। हम इसे इस प्रकार भी दिखा सकते हैं

$$\frac{3.00}{2.50}-1=0.20;$$

या प्रतिशत रूप में $0.20\times100=20$ प्रतिशत। अतः दिसंबर में प्याज की कीमतें अक्तूबर की तुलना में 20% अधिक हैं। इन दो महीनों की कीमतों के अनुपात

$$\frac{3.00}{2.50}=1.20$$

को **कीमत-सापेक्ष** (price relative) कहते हैं। कीमत सापेक्ष एक **शुद्ध संख्या** है।

कीमत सापेक्ष से '1' घटाकर सापेक्ष परिवर्तन को मापा जाता है। किसी एक वस्तु की कीमत सापेक्ष को उस वस्तु का सूचकांक कहते हैं।

एक दूसरा उदाहरण लें। हम दो वस्तुओं जैसे प्याज और कपड़े को लेते हैं। अक्तूबर 2001 में तथा दिसंबर 2001 में उनकी कीमतें निम्नलिखित थीं —

	अक्तूबर 2001	दिसंबर 2001
प्याज	2 रु. 50 पैसे प्रति कि.ग्रा.	3 रु. प्रति कि.ग्रा.
कपड़े	15 रु. प्रति मीटर	15 रु. 50 पैसे प्रति मीटर

दोनों ही स्थितियों में, दिसंबर में कीमतें अक्तूबर की तुलना में 50 पैसे अधिक थीं। परंतु ध्यान देने योग्य बात यह है कि कपड़े की कीमत **15 रु. प्रति**

मीटर में 50 पैसे प्रतिमीटर की वृद्धि की तुलना में प्याज की कीमत **2 रु. 50 पैसे प्रति कि.ग्रा.** में 50 पैसे प्रति कि.ग्रा. की वृद्धि अधिक महत्त्वपूर्ण है।

यदि हम कीमत में सापेक्ष परिवर्तन को निम्नलिखित विधि से मापें —

$$\frac{3.00-2.50}{2.50}\times 100 = 20$$

तो हम देखते हैं कि दिसंबर में प्याज की कीमत में अक्तूबर की तुलना में 20 प्रतिशत की वृद्धि हुई है, जबकि कपड़े की कीमत में सापेक्ष परिवर्तन को इस विधि से माप कर —

$$\frac{15.50-15.00}{15.00}\times 100=3.33$$

हम देखते हैं कि दिसंबर में कपड़े की कीमत में अक्तूबर की तुलना में 3.33 प्रतिशत की वृद्धि हुई है।

कीमत में वास्तविक संख्यात्मक अंतर महत्त्वपूर्ण नहीं है। महत्त्वपूर्ण यह है कि **मूल कीमतों की तुलना में** वर्तमान कीमतों में कितना परिवर्तन हुआ है?

ऊपर दिए गए उदाहरण में दो वस्तुएँ ली गई हैं। देखना यह है कि क्या हम प्याज तथा कपड़े की कीमतों में परिवर्तन का एक **संयुक्त सूचकांक** (Composite Index) निकाल सकतें हैं? कीमतों में निरपेक्ष अतंर को एक साथ नहीं मिलाया जा सकता है क्योंकि उनके मापन की इकाइयों में अंतर है। परंतु उनकी सापेक्ष कीमतें

$$\frac{3.00}{2.50} \text{ तथा } \frac{15.50}{15.00}$$

शुद्ध संख्याएँ हैं, अतः हम इन्हें सार्थक रूप से एक साथ मिला सकते हैं।

अब हम **सापेक्ष मूल्यों का समांतर** माध्य निकालेंगे:

$$\frac{1}{2}\left(\frac{3.00}{2.50} + \frac{15.50}{15.00}\right) = 1.12$$

जो प्याज तथा कपड़े की कीमतों का **संयुक्त सूचकांक** है। इस संयुक्त सूचकांक में से 1 रु. को घटा कर (1.12-1.00=) 0.12 प्राप्त होता है जो दोनों वस्तुओं की कीमतों में सापेक्ष परिवर्तन को एक साथ दिखाता है। प्रतिशत के रुप में दोनों वस्तुओं के **संयुक्त मूल्यों** में $0.12 \times 100=12$ प्रतिशत की वृद्धि हुई है।

यदि दो से अधिक वस्तुएँ दी गई हैं तो हम सापेक्ष कीमतों का समांतर माध्य निकालकर संयुक्त सूचकांक प्राप्त कर सकते हैं।

परंपरानुसार, सूचकांक प्रतिशत के रूप में व्यक्त किए जाते हैं।

2. उद्देश्य

कीमतों के सूचकांक वस्तुओं के समूह की कीमतों में सापेक्ष परिवर्तन को मापते हैं। इसी प्रकार, हम एक सूचकांक तैयार कर सकते हैं जो उत्पादन या उपभोग की गई विभिन्न वस्तुओं की 'मात्राओं' में सापेक्ष परिवर्तन को माप सकता है जिसे 'मात्रा सूचकांक' (Quantity Index) कहा जाता है।

सूचकांक का निर्माण करने के क्रम में कुछ प्रश्न महत्त्वपूर्ण हैं। उदाहरण के लिए,

i. किन वस्तुओं का समावेश होना चाहिए, तथा
ii. किस प्रकार की कीमतों (खुदरा या थोक) का प्रयोग होना चाहिए, आदि।

इन प्रश्नों के उत्तर उन उद्देश्यों पर निर्भर हैं, जिनके लिए सूचकांकों का निर्माण किया गया है। मान लें कि हमें सामान्य मूल्य-स्तर में परिवर्तन को मापना है। इसके लिए औद्योगिक वस्तुओं, कृषि संबंधी वस्तुओं तथा अन्य वस्तुओं की थोक कीमतों की आवश्यकता होती है। यदि हम निर्वाह-व्यय में परिवर्तन को मापना चाहें तो इसके लिए हमें उपभोक्ता वस्तुओं व सेवाओं की खुदरा कीमतों की जानकारी चाहिए। इसी प्रकार, कृषि तथा

औद्योगिक उत्पादन में परिवर्तन को मापने के लिए हमे कई प्रकार की सूचनाओं की आवश्यकता होती है।

3. आधार वर्ष का चुनाव

ऊपर दिए गए उदाहरणों में हमने अक्तूबर 2001 की कीमतों की तुलना में दिसंबर 2001 की सापेक्ष कीमतों में परिवर्तन पर विचार किया था। अतः अक्तूबर को **आधार अवधि** तथा दिसंबर को **वर्तमान अवधि** कहते हैं।

आधार अवधि का चुनाव (जो वस्तुतः संदर्भ-अवधि है) निम्नलिखित बातों पर निर्भर करता है :

क. **आधार-अवधि एक सामान्य अवधि होनी चाहिए।** असामान्य अवधियों जैसे युद्ध, अकाल या बाढ़ वाली अवधियों को आधार-अवधि नहीं चुना जाना चाहिए।

ख. **आधार-अवधि न तो बहुत छोटी हो और न ही बहुत लंबी।** एक कार्यकारी नियम के अनुसार यह न तो एक महीने से कम और न ही एक वर्ष से अधिक होनी चाहिए।

ग. **यह ऐसी अवधि होनी चाहिए जिसके लिए वास्तविक आँकड़े उपलब्ध हों।**

घ. **आधार-अवधि के रूप में बहुत पहले का समय नहीं लिया जाना चाहिए।** आधार-अवधि संदर्भ-अवधि मानी जाती है जिसके साथ वर्तमान अवधि की तुलना की जाती है। यदि यह बहुत पहले की है तो तुलना का कोई अर्थ नहीं रह जाता, क्योंकि समय के साथ पसंद बदलती है तथा वस्तुओं का चलन भी समाप्त हो जाता है। ऐसी स्थिति में पुरानी आधार-अवधि को बदलकर निकट की किसी अवधि को आधार बना लेना चाहिए।

4. वस्तुओं का चुनाव

सूचकांक के निर्माण के लिए कैसी और कितनी वस्तुएँ चुनी जानी चाहिए, यह इस पर निर्भर करता है कि सूचकांक का प्रयोग किस उद्देश्य के लिए किया जाता है।

किसी भी स्थिति में सभी वस्तुओं को शामिल करना संभव नहीं है। अतः सबसे पहले हमें उस समूह की पहचान करनी चाहिए जिसके लिए सूचकांक का निर्माण किया जा रहा हो।

यह आवश्यक है कि चुनी गई वस्तुएँ उस **समूह का प्रतिनिधित्व** करती हों।

उदाहरण स्वरूप, एक सामान्य कीमत-सूचकांक के निर्माण के लिए, हम कुछ प्रमुख औद्योगिक व कृषि वस्तुओं तथा अन्य वस्तुओं व सेवाओं की थोक कीमतों को शामिल करते हैं। इसी प्रकार उपभोक्ता कीमत-सूचकांक के निर्माण के लिए, हमें उन सभी वस्तुओं की खुदरा कीमतों को शामिल करना चाहिए जो परिवार का बजट बनाने में महत्त्वपूर्ण हों।

5. सूचकांकों का निर्माण (Construction of Index Numbers)

संकेत चिह्न (Notation)

'0' अंक का प्रयोग आधार-अवधि के लिए तथा '1' का वर्तमान अवधि के लिए किया जाता है। अतः p_{i1}, वर्तमान अवधि में i -वीं वस्तु की कीमत है तथा p_{io} आधार अवधि में इसकी कीमत है।

p_{o1} आधार अवधि 'o' की तुलना में वर्तमान अवधि '1' में कीमतों के सूचकांक को दर्शाता है। इसी प्रकार Q_{01} आधार अवधि 'o' की तुलना में, वर्तमान अवधि '1' में मात्राओं के सूचकांक को दर्शाता है।

मूल्य सूचकांक V_{01} को भी हम इसी प्रकार परिभाषित कर सकते हैं।

सरल सामुदायिक सूचकांक (Simple Aggregative Index)

एक सरल सामुदायिक कीमत सूचकांक को इस प्रकार दिखाते हैं

$$p_{01} = \frac{\sum p_{i1}}{\sum p_{io}} \times 100;$$

जो वर्तमान अवधि '1' में वस्तुओं की कीमतों का योग है तथा जिसे आधार अवधि '0' में वस्तुओं की कीमतों के योग के प्रतिशत के रूप में व्यक्त किया जाता है।

विभिन्न वस्तुओं की कीमतों के मापन की इकाइयाँ सामान्य रूप से समान नहीं होती हैं (रु. प्रति कि.ग्रा., रु. प्रति मीटर, रु. प्रति लीटर आदि)। अतः कीमतों के सीधे योग का प्रश्न ही नहीं उठता। इसलिए कीमतों या मात्राओं के सरल सामुदायिक सूचकांक की उपयोगिता सीमित है।

अभारित सूचकांक (Unweighted Index Number)

जैसा पहले बताया जा चुका है, एक या एक से अधिक वस्तुओं के लिए सूचकांक का निर्माण किया जा सकता है। यदि हमारे पास सिर्फ एक वस्तु हो तो सूचकांक आधार-अवधि की तुलना में वर्तमान अवधि में वस्तु की कीमत के अनुपात के रूप में सीधे दिखाया जा सकता है। इसे सामान्यतया प्रतिशत रूप में व्यक्त किया जाता है।

यदि वस्तुओं की संख्या एक से अधिक हो तो सूचकांक कीमत सापेक्षों के समांतर माध्य के रूप में प्रतिशत की तरह व्यक्त किया जा सकता है।

मान लें कि n वस्तुओं की कीमतें, वर्तमान अवधि '1' में, $p_{11}, p_{21} \cdots\cdots p_{n.1}$ हैं तथा आधार अवधि में $p_{10}, p_{20} \cdots\cdots p_{n0}$ हैं। ऐसी स्थिति में n वस्तुओं के कीमत-सापेक्ष इस प्रकार होंगे

$$\frac{p_{11}}{p_{10}}, \frac{p_{21}}{p_{20}}, \cdots\cdots, \frac{p_{n1}}{p_{n0}}$$

अभारित सूचकांकों को समांतर माध्य के रूप में इस प्रकार दिखाते हैं

$$\frac{1}{2}\left(\frac{p_{11}}{p_{10}} + \frac{p_{21}}{p_{20}} + \ldots\ldots + \frac{p_{n1}}{p_{n0}}\right)$$

या संक्षेप में $\frac{1}{n}\sum \frac{p_{i1}}{p_{io}}$

अभारित सूचकांक का दोष है कि यह सभी वस्तुओं को समान भार (महत्त्व) देता है, जबकि कुछ वस्तुएँ दूसरी वस्तुओं की तुलना में अधिक महत्त्वपूर्ण हो सकती हैं।

भारित सूचकांक (Weighted Index Numbers)

किसी संयुक्त सूचकांक के अंतर्गत भिन्न-भिन्न वस्तुओं को पर्याप्त महत्त्व देने के लिए, हम उन्हें उपयुक्त भार प्रदान करते हैं। भारित सूचकांक मूल्य सापेक्षों का भारित समांतर माध्य है जिसकी परिभाषा नीचे दी गई है :

$$p_{01} = \sum wi \frac{p_{i1}}{p_{io}}$$

जहाँ भार $w_1, w_2, \ldots w_n$ ऐसे दिए गए हैं ताकि

$$\sum w_i = 1$$

भिन्न-भिन्न वस्तुओं को भार प्रदान करने का सबसे आसान तरीका कुल व्यय में से उन वस्तुओं पर किए जाने वाले व्यय के अनुपात (या प्रतिशत) के रूप में भार प्रदान करना है।

उदाहरण

नीचे दिए गए आँकड़ों से कीमतों के 'भारित' तथा 'अभारित' सूचकांक का परिकलन करें :

वस्तु	कीमत आधार अवधि	भार वर्तमान अवधि	कीमत सापेक्ष	p_{i1}/p_{io}
गेहूँ	10 रु. प्रति कि.ग्रा.	15 रु. प्रति कि.ग्रा.	30%	1.50
चावल	15 रु. प्रति कि.ग्रा.	25 रु. प्रति कि.ग्रा.	40%	1.67
नमक	2 रु. प्रति कि.ग्रा.	2.50 रु. प्रति कि.ग्रा.	2%	1.25
घी	40 रु. प्रति कि.ग्रा.	60 रु. प्रति कि.ग्रा.	5%	1.50
दूध	12 रु. प्रति लीटर	15 रु. प्रति लीटर	20%	1.25
कपड़ा	40 रु. प्रति मीटर	60 रु. प्रति मीटर	3%	1.5
			100	**8.67**

कीमतों का अभारित सूचकांक है

$$\frac{1}{6}\sum\frac{p_{i1}}{p_{io}}=\frac{8.67}{6}=1.45$$

तथा **कीमतों का भारित सूचकांक है**

$$\frac{\sum wi\left(\frac{p_{i1}}{p_{io}}\right)}{\sum wi}=$$

$$\frac{(30\times1.50+40\times1.67+2\times1.25+5\times1.5+20\times1.25+3\times1.5}{100}=1.51$$

प्रतिशत के रूप में व्यक्त करने पर अभारित माध्य 145 है तथा भारित माध्य 151 है।

भारों का चुनाव (Selection of Weights)

विभिन्न वस्तुओं को दिए गए भार किसी समष्टि में उनके सापेक्ष महत्त्व को दिखाते हैं। ये भार कुल व्यय में विभिन्न वस्तुओं के 'मूल्य-भाग' कहे जाते हैं।

मान लें कि $p_{10}, p_{20},..........p_{no}$ आधार वर्ष की कीमतें हैं तथा $q_{10}, q_{20},..........q_{no}$ आधार वर्ष में उपभोग की गई विभिन्न वस्तुओं की तदनुरूप मात्राएँ हैं। अतः $p_{10}q_{10}, p_{20}q_{20},..........p_{no}q_{no}$ आधार अवधि में उपभोग की गई विभिन्न वस्तुओं के मूल्य हैं। आधार-अवधि का कुल व्यय है $\sum p_{io}q_{io}$। आधार अवधि में विभिन्न वस्तुओं के 'मूल्य भाग' निम्नलिखित हैं :

$$W_1=\frac{p_{10}q_{10}}{\sum p_{io}q_{io}}, W_2=\frac{p_{10}q_{10}}{\sum p_{io}q_{io}}.....W_n=\frac{p_{no}q_{no}}{\sum p_{io}q_{io}}$$

इन मूल्य भागों को भार मानते हुए, कीमत सापेक्ष का एक भारित औसत निम्नलिखित होगा :

$$p_{o1}=\sum wi\left(\frac{p_{i1}}{p_{io}}\right)$$

$$=\sum\frac{p_{io}q_{io}}{\sum p_{io}q_{io}}\left(\frac{p_{i1}}{q_{io}}\right)$$

$$=\frac{\sum p_{i1}q_{io}}{\sum p_{io}q_{io}}$$

इसे **लेस्पेयर सूचकांक** कहते हैं। इसके स्थान पर, हम आधार-अवधि की कीमतों पर वर्तमान अवधि में उपभोग की गई विभिन्न वस्तुओं के 'मूल्य भागों' का चुनाव कर सकते हैं। ऐसी स्थिति में भार निम्नलिखित होंगेः

$$W'_1=\frac{p_{10}q_{10}}{\sum p_{io}q_{io}}, W'_2=\frac{p_{20}q_{20}}{\sum p_{io}q_{io}}.....W'_n=\frac{p_{no}q_{no}}{\sum p_{io}q_{io}}$$

अतः सूचकांक निम्नलिखित होगा :

$$p_{o1}=\sum w'_i\left(\frac{p_{i1}}{p_{io}}\right)$$

$$=\sum\frac{p_{io}q_{i1}}{\sum p_{io}q_{i1}}\left(\frac{p_{i1}}{q_{io}}\right)$$

$$=\sum \frac{p_{i1}q_{i1}}{\sum p_{io}q_{i1}}$$

इसे **पाशे सूचकांक** कहते हैं।

6. प्रयोग में आने वाले कुछ महत्त्वपूर्ण सूचकांक

निम्नलिखित सूचकांक सरकार द्वारा नियमित प्रकाशित किए जाते हैं तथा नीतिगत उद्देश्यों के लिए प्रयोग किए जाते हैं :

i. उपभोक्ता कीमत सूचकांक (Consumer Price Index)

उपभोक्ता कीमत सूचकांक को **निर्वाह सूचकांक** भी कहते हैं। यह उन 'खुदरा कीमतों' में औसत परिवर्तन को मापता है जिन पर किसी निश्चित समय-बिंदु पर उपभोक्ता वस्तुओं व सेवाओं को खरीदते हैं।

चूँकि वस्तुओं की संख्या अधिक होती है अतः यह आवश्यक है कि उपभोक्ताओं के एक विशेष समूह का चुनाव किया जाए तथा उन उपभोक्ताओं के द्वारा उपभोग की गई वस्तुओं को ही चुना जाए।

उपभोक्ताओं के मूख्य समूह, जिनके लिए भारत में उपभोक्ता कीमत-सूचकांकों का निर्माण होता है, नीचे दिए जा रहे हैं:

क. औद्योगिक श्रमिक

ख. शारीरिक श्रम न करने वाले शहरी श्रमिक

ग. खेती करने वाले श्रमिक

औद्योगिक श्रमिकों के लिए उपभोक्ता कीमत सूचकांक सबसे अधिक प्रचलित सूचकांक है। इसकी रचना प्रतिवर्ष **1993-94 को आधार-वर्ष** मानकर की जाती है।

एक निश्चित अवधि में किसी विशेष समूह के परिवारों के, अपने कुल पारिवारिक व्यय में से विभिन्न मदों पर हुए व्यय की जानकारी प्राप्त करने के लिए एक प्रतिदर्श सर्वेक्षण किया जाता है।

इसके बाद पूरे समूह के लिए **एक प्रतिनिधि** पारिवारिक बजट तैयार किया जाता है। इसका प्रयोग उपभोक्ता कीमत सूचकांक के निर्माण के लिए वस्तुओं व सेवाओं तथा उनके सापेक्ष भारों (कुल पारिवारिक व्यय में उनके सापेक्ष भागों) के चुनाव के लिए किया जाता है।

उपभोक्ता मूल्य सूचकांकों के उपयोग

खुदरा कीमतों में वृद्धि के कारण उपभोक्ता मूल्य सूचकांक में वृद्धि होती है। मूद्रा की क्रय-शक्ति में कमी हो जाती है तथा वास्तविक मजदूरी घट जाती है।

क. उपभोक्ता कीमत सूचकांक का प्रयोग मुद्रा की क्रय शक्ति तथा वास्तविक मजदूरी निर्धारित करने के लिए किया जाता है।

मान लें कि उपभोक्ता मूल्य सूचकांक 1998-99 में 250 था, जबकि यह आधार अवधि 1993-94 में 100 था। अतः 1998-99 में एक रुपया 1993-94 के $\left(\frac{100}{250}=0.40\right)$ 40 पैसे के बराबर होगा। यह 1998-99 में रु. की क्रय शक्ति में 60 पैसे की गिरावट को दर्शाती है।

यदि उपभोक्ता की मुद्रा मजदूरी 1998-99 में 800 रु. प्रति माह थी तो उसकी वास्तविक मजदूरी (1993-94 की कीमतों के अनुसार)

$$800\times\frac{100}{250}=320 \text{ रु. प्रति माह होगी।}$$

उपभोक्ता मूल्य सूचकांक को **मूल्य का अपस्फीतिकारक** (Price Deflator of Income) कहते हैं।

ख. जब उपभोक्ता मूल्य सूचकांक किसी निश्चित स्तर से ऊपर जाता है तो सरकार अतिरिक्त धन देकर श्रमिकों की क्षतिपूर्ति करने का निश्चय करती है, जैसे महँगाई भत्ता। सहायता का परिमाण उपभोक्ता मूल्य सूचकांक में वृद्धि द्वारा निर्धारित होता है।

ग. यदि कुछ आवश्यक वस्तुओं (जैसे गेहूँ, चावल, चीनी, कपड़ा आदि) की कीमतें उनके अभाव के कारण बढ़ती हैं, तो सरकार इन्हें उचित दर की दूकानों या राशन द्वारा प्रदान करने का निश्चय कर सकती है।

ii. थोक मूल्य सूचकांक (Wholesale Price Index)

थोक मूल्य सूचकांक 'सामान्य मूल्य स्तर' में परिवर्तन को सूचित करता है। इसका **साप्ताहिक प्रकाशन** नियमित रूप से उद्योग-मंत्रालय, भारत सरकार के आर्थिक सलाहकार के कार्यालय से किया जाता है।

सरकार द्वारा प्रकाशित थोक मूल्य सूचकांक की नई श्रेणी में आधार-अवधि 1993-94 को लिया गया है, जबकि पुरानी श्रेणी में आधार-अवधि 1981-82 था। इसके अंतर्गत वस्तुओं की एक बड़ी संख्या शामिल है, जो निम्नलिखित प्रमुख समूहों में बाँटी गई है :

क. प्राथमिक वस्तुएँ जैसे खाद्य व गैर-खाद्य मदें,

ख. ईंधन, ऊर्जा, प्रकाश, लुब्रिकेंट, तथा

ग. विनिर्मित वस्तुएँ

इनमें से प्रत्येक समूह को उपयुक्त भार प्रदान किए जाते हैं।

थोक मूल्य सूचकांक की उपयोगिताएँ

क. थोक मूल्य सूचकांकों की काल-श्रेणियों का प्रयोग भविष्य के मूल्यों का पूर्वानुमान लगाने के लिए किया जा सकता है।

ख. चूँकि कीमतें माँग व पूर्ति दोनों को प्रभावित करती हैं, अतः एक उपयुक्त निदर्श का प्रयोग करके हम भविष्य में माँग व पूर्ति की स्थितियों का अनुमान लगा सकते हैं।

ग. थोक मूल्य सूचकांक का प्रयोग **मुद्रा-स्फीति की दर** जानने के लिए किया जा सकता है। मान लें t वें तथा (t-1) वें सप्ताह में थोक मूल्य सूचकांक क्रमशः X_t तथा X_{t-1} हैं, अतः **साप्ताहिक मुद्रस्फीति दर** का परिकलन हम इस प्रकार कर सकते हैं —

$$\frac{X_t - X_{t-1}}{X_{t-1}} \times 100$$

थोक मूल्य सूचकांकों के **वार्षिक औसतों** का प्रयोग करके वार्षिक मुद्रास्फीति दर का परिकलन किया जाता है।

उदाहरण के लिए, 1997-98 तथा 1998-99 में थोक मूल्य सूचकांकों के वार्षिक औसत क्रमशः 134.4 तथा 142.4 थे। अतः इस अवधि में वार्षिक मुद्रास्फीति दर निम्नलिखित थी :

$$\frac{142.4 - 134.4}{134.4} \times 100 = 5.95\%$$

विभिन्न वस्तुओं या वस्तु-समूहों के लिए मुद्रा-स्फीति दरों का आकलन भी नीति-निर्धारण के उद्देश्य से किया जाता जा सकता है।

घ. थोक मूल्य सूचकांक का प्रयोग राष्ट्रीय आय, पूँजी निर्माण आदि पर मूल्यों में परिवर्तन के प्रभाव को दूर करने के लिए किया जाता है। किसी वर्ष में उत्पादित वस्तुओं व सेवाओं के मूल्यों को राष्ट्रीय आय कहते हैं। यदि हम वस्तुओं व सेवाओं के मूल्यों का परिकलन उस वर्ष के प्रचलित मूल्यों के अनुसार करें तो हमें **वर्तमान मूल्यों पर राष्ट्रीय आय** प्राप्त होती है।

वर्तमान मूल्यों पर राष्ट्रीय आय में वृद्धि निम्नलिखित कारणों से हो सकती है :

क. सामान्य मूल्य-स्तर में वृद्धि, या

ख. वास्तविक उत्पादन में वृद्धि।

वास्तविक उत्पादन को निर्धारित करने के लिए, हमें कीमतों में परिवर्तन के प्रभाव को दूर रखना चाहिए। इसे निम्नलिखित उदाहरण से स्पष्ट करें।

उदाहरण

मान लें कि 1997-98 में वर्तमान मूल्यों पर कुल उत्पादन 1000 करोड़ रु. था तथा पिछले वर्ष 1996-97 में यह 885 करोड़ रु. था। 1997-98 में थोक मूल्य सूचकांक 134.4 तथा 1996-97 में 128 था। 1996-97 के मूल्यों के आधार पर 1997-98 में उत्पादन का परिकलन इस प्रकार करें :

$$\frac{128}{134.4} \times 1000 = 952$$

अतः वास्तविक उत्पादन में वृद्धि केवल 952-885=67 थी न कि 1000-885=115 ।

आधार-वर्ष के मूल्यों पर परिकलन किया गया वास्तविक उत्पादन $\frac{100}{134.4} \times 1000 = 744$ होगा।

iii. औद्योगिक उत्पादन का सूचकांक (Index of Industrial Production)

औद्योगिक उत्पादन का सूचकांक मात्रा सूचकांक कहलाता है। यह औद्योगिक उत्पादन में परिवर्तन की माप करता है।

इस सूचकांक का निर्माण भारत में वार्षिक रूप से किया जाता है जिसमें वर्तमान शृंखला के लिए 1993-94 को आधार वर्ष माना गया है।

जिन तीन प्रमुख वर्गों के लिए सूचकांकों का निर्माण किया जाता है, वे हैं :

क. खनन,

ख. विनिर्माण तथा

ग. ऊर्जा

इन तीनों को उपयुक्त भार प्रदान किए जाते हैं।

अभ्यास

1. (क) मूल्यों, तथा (ख) मात्राओं के सूचकांकों के निर्माण के उद्देश्य क्या हैं?
2. सूचकांक के निर्माण का एक सामान्य तरीका बताएँ तथा इसके प्रयोगों के बारे में बताएँ।
3. सूचकांकों के निर्माण में आधार-अवधि के चुनाव में किन बातों का ध्यान रखा जाता है?
4. सरल सामुदायिक मूल्य-सूचकांक का वर्णन करें। इसके दोष क्या हैं?
5. मूल्यों के 'भारित' तथा 'अभारित' सूचकांकों का वर्णन करें।
6. मूल्यों के भारित सूचकांकों के निर्माण में
   (क) भारों, तथा
   (ख) वस्तुओं के चुनाव में किन बातों का ध्यान रखा जाता है?

7. निम्नलिखित सारणी में एक साथ 1991 तथा 2001 में मुख्य खाद्य-सामग्रियों की कीमतें तथा 1991 में प्रत्येक परिवार के प्रति माह औसत व्यय दिए गए हैं :

मदें	इकाइयाँ	रुपयों में कीमतें		1991 में औसत व्यय/
		1991	2001	प्रति माह (रु. में)
चावल	रु.प्रति कि.ग्रा.	15.00	35.00	804
गेहूँ	रु.प्रति कि.ग्रा.	9.00	15.00	310
दालें	रु.प्रति कि.ग्रा.	20.00	25.00	245
दूध	रु.प्रति लीटर	7.00	15.00	115
तेल	रु.प्रति लीटर	30.00	65.00	110
मछली	रु.प्रति कि.ग्रा.	30.00	75.00	260
चाय	रु.प्रति कप	0.75	3.00	130

निम्नलिखित का परिकलन करें तथा परिणामों की तुलना करें :

क. सरल सामुदायिक मूल्य सूचकांक

ख. मूल्य सापेक्षों के अभारित सूचकांक

ग. मूल्य सापेक्षों के उपयुक्त भारित सूचकांक

8. खाद्यान्नों तथा गैर-खाद्यान्नों के उत्पादन के सूचकांक नीचे दिए गए हैं :

वर्ष	कृषि उत्पादन के सूचकांक	
	खाद्यान्न भार = 62.92	गैर-खाद्यान्न भार = 37.08
1993-94	150.2	169.4
1994-95	155.9	180.9
1995-96	146.1	185.4
1996-97	160.9	200.9
1997-98	155.7	180.6
1998-99	164.8	198.1

कुल कृषि उत्पादन का सूचकांक निकालें।

9. 1981- 82 (=100) को आधार मान कर प्रमुख फसलों को उपजाने वाले क्षेत्रों के सूचकांक नीचे दिए गए हैं :

(आधार : 1981 –82 =100)

वर्ष	खाद्यान्न	गैर-खाद्यान्न
1990-91	100.7	120
1993-94	96.7	127.3
1994-95	97.6	126.2
1995-96	95.3	131.7
1996-97	97.4	134.6
1997-98	97.6	133.5
1998-99	98.8	135.4

खाद्यान्नों में निम्नलिखित सम्मिलित हैं :

क. अनाज (चावल, गेहूँ, मोटे अनाज), तथा

ख. दालें

गैर-खाद्यान्नों में निम्नलिखित सम्मिलित हैं :

क. तिलहन (मूँगफली, रेपसीड तथा सरसों)

ख. रेशेदार फसलें (कपास, जूट आदि)

ग. वृक्षों वाली फसलें (चाय, काफी, रबर), तथा

घ. अन्य (गन्ना, तम्बाकू, आलू)

आँकड़ों की व्याख्या करें। इनसे आप क्या निष्कर्ष निकालते हैं? एक संक्षिप्त टिप्पणी लिखें।

अध्याय - 10

अर्थशास्त्र में सांख्यिकीय विधियों के प्रयोग की परियोजना

यदि हम किसी सामाजिक-आर्थिक समस्या के अध्ययन के लिए कोई परियोजना बनाते हैं तो हमें चरों से संबद्ध आँकड़ों को संग्रहीत करने तथा उनका विश्लेषण करने की आवश्यकता होती है। आँकड़े 'द्वितीयक' हो सकते हैं (जो प्रकाशित स्रोतों से लिए गए हों) या 'प्राथमिक' (जो अन्वेषणकर्ता के द्वारा स्वयं क्षेत्रीय सर्वेक्षण करके एकत्र किए गए हों)। हमने अध्याय 2 में प्राथमिक आँकड़ों को संग्रहीत करने की विधियों (सर्वेक्षण तथा प्रतिदर्श विधियों) तथा संबंधित समस्याओं का वर्णन किया है। अब हम निम्नलिखित बिंदुओं पर ध्यान देंगे :

क. अध्ययन के उद्देश्य तथा लक्ष्य को स्पष्ट रूप से व्यक्त किया जाना चाहिए

उदाहरण के लिए, विक्रेताओं द्वारा किए जाने वाले अपराधों की जाँच (Assessment) करने के लिए हम 'परिवारों में उपभोक्ता-अधिकारों के प्रति जागरूकता' पर एक अध्ययन की योजना बना सकते हैं या उत्पादन प्रक्रिया में सुधार के लिए हम 'उद्यमों की उत्पादन क्षमता के प्रति जागरूकता' पर एक अध्ययन की योजना बना सकते हैं।

ख. समष्टि का निर्धारण किया जाना चाहिए

यदि परिवारों में उपभोक्ता अधिकारों के प्रति जागरूकता की जानकारी के लिए क्षेत्रीय सर्वेक्षण करना हो तो हम सबसे पहले परिवारों के उस समूह की पहचान करेंगे जिन पर विचार करना हो। कुछ परिवार बहुत धनी होते हैं तथा कुछ बहुत गरीब। कुछ मध्यवर्गीय परिवार भी होते हैं। इन वर्गों में खरीदी जाने वाली वस्तुओं की गुणवत्ता से संबंधित जानकारी में बहुत अंतर होता है। शहरी व ग्रामीण तथा शिक्षित व अशिक्षित परिवारों की प्रतिक्रियाएँ भिन्न-भिन्न होती हैं। इसी प्रकार यदि हम उद्यमों की 'उत्पादन की क्षमता की जागरूकता' का अध्ययन करना चाहते हों, तो हमें बड़े तथा छोटे, दोनों प्रकार के उद्यमों का अध्ययन करना होता है, जिनकी उत्पादकता के प्रति जागरूकता में काफी अंतर होता है।

ग. विश्लेषण की विधि

यदि हम प्रकाशित स्रोतों से द्वितीयक आँकड़ों का प्रयोग कर रहे हों तो हमें उन चरों की परिभाषा का ध्यान रखना चाहिए जिनके लिए आँकड़े एकत्र करने हों। हमें इसकी जानकारी लेनी चाहिए कि आँकड़े किस प्रकार एकत्र किए गए हैं। उदाहरण के लिए, यदि ये सूचकांक हैं तो इनके निर्माण की विधि (भार, विस्तार आदि) क्या थी? यदि ये आँकड़े फसल के उत्पादन से संबद्ध हैं तो इनके मूल्य वर्तमान कीमतों पर लिए गए हैं या स्थायी कीमतों पर।

प्रकाशित आँकड़ों का प्रयोग सावधानीपूर्वक किया जाना चाहिए, अन्यथा हम गलत निष्कर्ष पर पहुँच सकते हैं।

यदि हम परिवारों के बीच उपभोक्ता जागरूकता के अध्ययन के लिए क्षेत्रीय सर्वेक्षण की योजना बना रहे हों तो हम यह देखते हैं कि समष्टि में सजातीयता नहीं हो (जैसा ऊपर दिया गया है) । अतः या तो हम सजातीय समूह लें सकते हैं या स्तरित यादृच्छिक प्रतिचयन विधि का प्रयोग कर सकते हैं। स्तरित यादृच्छिक प्रतिचयन विधि का वर्णन अध्याय 2 में किया गया है।

घ. प्रश्नावली

प्रश्नावली को सावधानीपूर्वक तैयार किया जाना चाहिए, जैसा कि परिशिष्ट 'ख' में सुझाया गया है। प्रश्नावली को कई भागों में बनाया जाना चाहिये, जैसा परिशिष्ट 'ख' में दिया गया है।

जैसा अध्याय 2 में बताया गया है, उत्तरदाताओं के लिए **साक्षात्कार विधि** या **डाक प्रश्नावली विधि** का प्रयोग किया जा सकता है। दोनों ही विधियों के अपने-अपने गुण तथा दोष हैं। (देखिए, अध्याय 2)।

ङ. आँकड़ों का विश्लेषण

हम उन परिवारों का अनुपात निकाल सकते है :

क. जो वस्तुएँ खरीदने से पहले उनकी जाँच करते हों।

ख. जिन्होंने विक्रेताओं को दोषपूर्ण व मिलावटयुक्त वस्तुओं की सूचना दी हो।

ग. जिनकी शिकायतों की सुनवाई विक्रेताओं द्वारा संतोषजनक ढंग से की गई हो।

घ. जिन्होंने उपभोक्ता-अदालतों में रिपोर्टें दर्ज की हों।

क्या ये अनुपात भिन्न-भिन्न स्तरों के बीच महत्त्वपूर्ण रूप से बदलते हैं?

क्या आप यह परिणाम निकालते हैं कि शिक्षित परिवारों में अधिक उपभोक्ता जागरूकता है?

परिशिष्ट 'क'

यादृच्छिक संख्या सारणी (i)

03 47 43 73 86	36 96 47 36 61	46 98 63 71 62	33 26 16 80 45	60 11 14 10 95
97 74 24 67 62	42 81 14 57 20	42 53 32 37 32	27 07 36 07 51	24 51 79 89 73
16 76 62 27 66	56 50 26 71 07	32 90 79 78 53	13 55 38 58 59	88 97 54 14 10
12 56 85 99 26	96 96 68 27 31	05 03 72 93 15	57 12 10 14 21	88 26 49 81 76
55 59 56 35 64	38 54 82 46 22	31 62 43 09 90	06 18 44 32 53	23 83 01 30 30
16 22 77 94 39	49 54 43 54 82	17 37 93 23 78	87 35 20 96 43	84 26 34 91 64
84 42 17 53 31	57 24 55 06 88	77 04 74 47 67	21 76 33 50 25	83 92 12 06 76
63 01 63 78 59	16 95 55 67 19	98 10 50 71 75	12 86 73 58 07	44 39 52 38 79
33 21 12 34 29	78 64 56 07 82	52 42 07 44 38	15 51 00 13 42	99 66 02 79 54
57 60 86 32 44	09 47 27 96 54	49 17 46 09 62	90 52 84 77 27	08 02 73 43 28
18 18 07 92 46	44 17 16 58 09	79 83 86 19 62	06 76 50 03 10	55 23 64 05 05
26 62 38 97 75	84 16 07 44 99	83 11 46 32 24	20 14 85 88 45	10 93 72 88 71
23 42 40 64 74	82 97 77 77 81	07 45 32 14 08	32 98 94 07 72	93 85 79 10 75
52 36 28 19 95	50 92 26 11 97	00 56 76 31 38	80 22 02 53 53	86 60 42 04 53
37 85 94 35 12	83 39 50 08 30	42 34 07 96 88	54 42 06 87 98	35 85 29 48 39
70 29 17 12 13	40 33 20 38 26	13 89 51 03 74	17 76 37 13 04	07 74 21 19 30
56 62 18 37 35	96 83 50 87 75	97 12 25 93 47	70 33 24 03 54	97 77 46 44 80
99 49 57 22 77	88 42 95 45 72	16 64 36 16 00	04 43 18 66 79	94 77 24 21 90
16 08 15 04 72	33 27 14 34 09	45 59 34 68 49	12 72 07 34 45	99 27 72 95 14
31 16 93 32 43	50 27 89 87 19	20 15 37 00 49	52 85 66 60 44	38 68 88 11 80
68 34 30 13 70	55 74 30 77 40	44 22 78 84 26	04 33 46 09 52	68 07 97 06 57
74 57 25 65 76	59 29 97 68 60	71 91 38 67 54	13 58 18 24 76	15 54 55 95 52
27 42 37 86 53	48 55 90 65 72	96 57 69 36 10	96 46 92 42 45	97 60 49 04 91
00 39 68 29 61	66 37 32 20 30	77 84 57 03 29	10 45 65 04 26	11 04 96 67 24
29 94 98 94 24	68 49 69 10 82	53 75 91 93 30	34 25 20 57 27	40 48 73 51 92
16 90 82 66 59	83 62 64 11 12	67 19 00 71 74	60 47 21 29 68	02 02 37 03 31
11 27 94 75 06	06 09 19 74 66	02 94 37 34 02	76 70 90 30 86	38 45 94 30 38
35 24 10 16 20	33 32 51 26 38	79 78 45 04 91	16 92 53 56 16	02 75 50 95 98
38 23 16 86 38	42 38 97 01 50	87 75 66 81 41	40 01 74 91 62	48 51 84 08 32
31 96 25 91 47	96 44 33 49 13	34 86 82 53 91	0 52 43 48 85	27 55 26 89 62
66 67 40 67 14	64 05 71 95 86	11 05 65 09 68	76 83 20 37 90	57 16 00 11 66
14 90 84 45 11	75 73 88 05 90	52 27 41 14 86	22 98 12 22 08	07 52 74 95 80
68 05 51 18 00	33 96 02 75 19	07 60 62 93 55	59 33 82 43 90	49 37 38 44 59
20 46 78 73 90	97 51 40 14 02	04 02 33 31 08	39 54 16 49 36	47 95 93 13 30
64 19 58 97 79	15 06 15 93 20	01 90 10 75 06	40 78 78 89 62	02 67 74 17 33
05 26 93 70 60	22 35 85 15 13	92 03 51 59 77	59 56 78 06 83	52 91 05 70 74
07 97 10 88 23	09 98 42 99 64	61 71 62 99 15	06 51 29 16 93	58 05 77 09 51
68 71 86 85 85	54 87 66 47 54	73 32 08 11 12	44 95 92 63 16	29 56 24 29 48
26 99 61 65 53	58 37 78 80 70	42 10 50 67 42	32 17 55 85 74	94 44 67 16 94
14 65 52 68 75	87 59 36 22 41	26 78 63 06 55	13 08 27 01 50	15 29 39 39 43

परिशिष्ट 'क' (क्रमशः)

(ii)

17 53 77 58 71	71 41 61 50 72	12 41 94 96 26	44 95 27 36 99	02 96 74 30 83
90 26 59 21 19	23 52 23 33 12	96 93 02 18 39	07 02 18 36 07	25 99 32 70 23
41 23 52 55 99	31 04 49 69 96	10 47 48 45 88	13 41 43 89 20	97 17 14 49 17
60 20 50 81 69	31 99 73 68 68	35 81 33 03 76	24 30 12 48 60	18 99 10 72 34
91 25 38 05 90	94 58 28 41 36	45 37 59 03 09	90 35 57 29 12	82 62 54 65 60
34 50 57 74 37	98 80 33 00 91	09 77 93 19 82	74 94 80 04 04	45 07 31 66 49
85 22 04 39 43	73 81 53 94 79	33 62 46 86 28	08 31 54 46 31	53 94 13 38 47
09 79 13 77 48	73 82 97 22 21	05 03 27 24 83	72 89 44 05 60	35 80 39 94 88
88 75 80 18 14	22 95 75 42 49	39 32 82 22 49	02 48 07 70 37	16 04 61 67 87
90 96 23 70 00	39 00 03 06 90	55 85 78 38 36	94 37 30 69 32	90 89 00 76 33
53 74 23 99 67	61 32 28 69 84	94 62 67 86 24	98 33 41 19 95	47 53 53 38 09
63 38 06 86 54	99 00 65 26 94	02 82 90 23 07	79 62 67 80 60	75 91 12 81 19
35 30 58 21 46	06 72 17 10 94	25 21 31 75 96	49 28 24 00 49	55 65 79 78 07
63 43 36 82 69	65 51 18 37 88	61 38 44 12 45	32 92 85 88 65	54 34 81 85 35
98 25 37 55 26	01 91 82 81 46	74 71 12 94 97	24 02 71 37 07	03 92 18 66 75
02 63 21 17 69	71 50 80 89 56	38 15 70 11 48	43 40 45 86 98	00 83 26 91 03
64 55 22 21 82	48 22 28 06 00	61 54 13 43 91	82 78 12 23 29	06 66 24 12 27
85 07 26 13 89	01 10 07 82 04	59 63 69 36 03	69 11 15 83 80	13 29 54 19 28
58 54 16 24 15	51 54 44 82 00	62 61 65 04 69	38 18 65 18 97	85 72 13 49 21
34 85 27 84 87	61 48 64 56 26	90 18 48 13 26	37 70 15 42 57	65 65 80 39 07
03 92 18 27 46	57 99 16 96 56	30 33 72 85 22	84 64 38 56 98	99 01 30 98 64
62 95 30 27 59	37 75 41 66 48	86 97 80 61 45	23 53 04 01 63	45 76 08 64 27
08 45 93 15 22	60 21 75 46 91	98 77 27 85 42	28 88 61 08 84	69 62 03 42 73
07 08 55 18 40	45 44 75 13 90	24 94 96 61 02	57 55 66 83 15	73 42 37 11 61
01 85 89 95 66	51 10 19 34 88	15 84 97 19 75	12 76 39 43 78	64 63 91 08 25
72 84 71 14 35	19 11 58 49 26	50 11 17 17 76	86 31 57 20 18	95 60 78 46 75
88 78 28 16 84	13 52 53 94 53	75 45 69 30 96	73 89 65 70 31	99 17 43 48 76
45 17 75 65 57	28 40 19 72 12	25 12 74 75 67	60 40 60 81 19	24 62 01 61 16
96 76 28 12 54	22 01 11 94 25	71 96 16 16 88	68 64 36 74 45	19 59 50 88 92
43 31 67 72 30	24 02 94 08 63	38 32 36 66 02	69 36 38 25 39	48 03 45 15 22
50 44 66 44 21	66 06 58 05 62	68 15 54 35 02	42 35 48 96 32	14 52 41 52 48
22 66 22 15 86	26 63 75 41 99	58 42 36 72 24	58 37 52 18 51	03 37 18 39 11
96 24 40 14 51	23 22 30 88 57	95 67 47 29 83	94 69 40 06 07	18 16 36 78 86
31 73 91 61 19	60 20 72 93 48	98 57 07 23 69	65 95 39 69 58	56 80 30 19 44
78 60 73 99 84	43 89 94 36 45	56 69 47 07 41	90 22 91 07 12	78 35 34 08 72
84 37 90 61 56	70 10 23 98 05	85 11 34 76 60	76 48 45 34 60	01 64 18 39 96
36 67 10 08 23	98 93 35 08 86	99 29 76 29 81	33 34 91 58 93	63 14 52 32 52
07 28 59 07 48	89 64 58 89 75	83 85 62 27 89	30 14 78 56 27	86 63 59 80 02
10 15 83 87 60	79 24 31 66 56	21 48 24 06 93	91 98 94 05 49	01 47 59 38 00
55 19 68 97 65	03 73 52 16 56	00 53 55 90 27	33 42 29 38 87	22 13 88 83 34
53 81 29 13 39	35 01 20 71 34	62 33 74 82 14	53 73 19 09 03	56 54 29 56 93
51 86 32 68 92	33 98 74 66 99	40 14 71 94 58	45 94 19 38 81	14 44 99 81 07
35 91 70 29 13	80 03 54 07 27	96 94 78 32 66	50 95 52 74 33	13 80 55 62 54
37 71 67 95 13	20 02 44 95 94	64 85 04 05 72	01 32 90 76 14	53 89 74 60 41
93 66 13 83 27	92 79 64 64 72	28 54 96 53 84	48 14 52 98 94	56 07 93 89 30

परिशिष्ट 'क' (क्रमशः)

(iii)

02 96 08 45 65	13 05 00 41 84	93 07 54 72 59	21 45 57 09 77	19 48 56 27 44
49 83 43 48 35	82 88 33 69 96	72 36 04 19 76	47 45 15 18 60	82 11 08 95 97
84 60 71 62 46	40 80 81 30 37	34 39 23 05 38	25 15 35 71 30	88 12 57 21 77
18 17 30 88 71	44 91 14 88 47	89 23 30 63 15	56 34 20 47 89	99 82 93 24 98
79 69 10 61 78	71 32 76 95 62	87 00 22 58 40	92 54 01 75 25	43 11 71 99 31
75 93 36 57 83	56 20 14 82 11	74 21 97 90 65	96 42 68 63 86	74 54 13 26 94
38 30 92 29 03	06 28 81 39 38	62 25 06 84 63	61 29 08 93 67	04 32 92 08 09
51 29 50 10 34	31 57 75 95 80	51 97 02 74 77	76 15 48 49 44	18 55 63 77 09
21 31 38 86 24	37 79 81 53 74	73 24 16 10 33	52 83 90 94 76	70 47 14 54 36
29 01 23 87 88	58 02 39 37 67	42 10 14 20 92	16 55 23 42 45	54 96 09 11 06
95 33 95 22 00	18 74 72 00 18	38 79 58 69 32	81 76 80 26 92	82 80 84 25 39
90 84 60 79 80	24 36 59 87 38	82 07 53 89 35	96 35 23 79 18	05 98 90 07 35
46 40 62 98 82	54 97 20 56 95	15 74 80 08 32	16 46 70 50 80	67 72 16 42 79
20 31 89 03 43	38 46 82 68 72	32 14 82 99 70	80 60 47 18 97	63 49 30 21 30
71 59 73 05 50	08 22 23 71 77	91 01 93 20 49	82 96 59 26 94	66 39 67 98 60
22 17 68 65 84	68 95 23 92 35	87 02 22 57 51	61 09 43 95 06	58 24 82 03 47
19 36 27 59 46	13 79 93 37 55	39 77 32 77 09	85 52 05 30 62	47 83 51 62 74
16 77 23 02 77	09 61 87 25 21	28 06 24 25 93	16 71 13 59 78	23 05 47 47 25
78 43 76 71 61	20 44 90 32 64	97 67 63 99 61	46 38 03 93 22	69 81 21 99 21
03 28 28 26 08	73 37 32 04 05	69 30 16 09 05	88 69 58 28 99	35 07 44 75 47
93 22 53 64 39	07 10 63 76 35	87 03 04 79 88	08 13 13 85 51	55 34 57 72 69
78 76 58 54 74	92 38 70 96 92	52 06 79 79 45	82 63 18 27 44	69 66 92 19 09
23 68 35 26 00	99 53 93 61 28	52 70 05 48 34	56 65 05 61 86	90 92 10 70 80
15 39 25 70 99	93 86 52 77 65	15 33 59 05 28	22 87 26 07 47	86 96 98 29 06
58 71 96 30 24	18 46 23 34 27	85 13 99 24 44	49 18 09 79 49	74 16 32 23 02
57 35 27 33 72	24 53 63 94 09	41 10 76 47 91	44 04 95 49 66	39 60 04 59 81
48 50 86 54 48	22 06 34 72 52	82 21 15 65 20	33 29 94 71 11	15 91 29 12 03
61 96 48 95 03	07 16 39 33 66	98 56 10 56 79	77 21 30 27 12	90 49 22 23 62
36 93 89 41 26	29 70 83 63 51	99 74 20 52 36	87 09 41 15 09	98 60 16 03 03
18 87 00 42 31	57 90 12 02 07	23 47 37 17 31	54 08 01 88 63	39 41 88 92 10
88 56 53 27 59	33 35 72 67 47	77 34 55 45 70	08 18 27 38 90	16 95 86 70 75
09 72 95 84 29	49 41 31 06 70	42 38 06 45 18	64 84 73 31 65	52 53 37 97 15
12 96 88 17 31	65 19 69 02 83	60 75 86 90 68	24 64 19 35 51	56 61 87 39 12
85 94 57 24 16	92 09 84 38 76	22 00 27 69 85	29 81 94 78 70	21 94 47 90 12
38 64 43 59 98	98 77 87 68 07	91 51 67 62 44	40 98 05 93 78	23 32 65 41 18
53 44 09 42 72	00 41 86 79 79	68 47 22 00 20	35 55 31 51 51	00 83 63 22 55
40 76 66 26 84	57 99 99 90 37	36 63 32 08 58	37 40 13 68 97	87 64 81 07 83
02 17 79 18 05	12 51 52 57 02	22 07 90 47 03	28 14 11 30 79	20 69 22 40 98
95 17 82 06 53	31 51 10 96 46	92 06 88 07 77	56 11 50 81 69	40 23 72 51 39
35 76 22 42 92	96 11 83 44 80	34 68 35 48 77	33 42 40 90 60	73 96 53 97 86
26 29 13 56 41	85 47 04 66 08	34 72 57 59 13	82 43 80 46 15	38 26 61 70 04
77 80 20 75 82	72 82 32 99 90	63 95 73 76 63	89 73 44 99 05	48 67 26 43 18
46 40 66 44 52	91 36 74 43 53	30 82 13 54 00	78 45 63 98 35	55 03 36 67 68
37 56 08 18 09	77 53 84 46 47	31 91 18 95 58	24 16 74 11 53	44 10 13 85 57
61 65 61 68 66	37 27 47 39 19	84 83 70 07 48	53 21 40 06 71	95 06 79 88 54

परिशिष्ट 'क' (क्रमशः)

(iv)

93 43 69 64 07	34 18 04 52 35	56 27 09 24 86	61 85 53 83 45	19 90 70 99 00
21 96 60 12 99	11 20 99 45 18	48 13 93 55 34	18 37 79 49 90	65 97 38 20 46
95 20 47 97 97	27 37 83 28 71	00 06 41 41 74	45 89 09 39 84	51 67 11 52 49
97 86 21 78 73	10 65 81 92 59	58 76 17 14 97	04 76 62 16 17	17 95 70 45 80
69 92 06 34 13	59 71 74 17 32	27 55 10 24 19	23 71 82 13 74	63 52 52 01 41
04 31 17 21 56	33 73 99 19 87	26 72 39 27 67	53 77 57 68 93	60 61 97 22 61
61 06 98 03 91	87 14 77 43 96	43 00 65 98 50	45 60 33 01 07	98 99 46 50 47
85 93 85 86 88	72 87 08 62 40	16 06 10 89 20	23 21 34 74 97	76 38 03 29 63
21 74 32 47 45	73 96 07 94 52	09 65 90 77 47	25 76 16 19 33	53 05 70 53 30
15 69 53 82 80	79 96 23 53 10	65 39 07 16 29	45 33 02 43 70	02 87 40 41 45
02 89 08 04 49	20 21 14 68 86	87 63 93 95 17	11 29 01 95 80	35 14 97 35 33
87 18 15 89 79	85 43 01 72 73	08 61 74 51 69	89 74 39 82 15	94 51 33 41 67
98 83 71 94 22	59 97 50 99 52	08 52 85 08 40	87 80 61 65 31	91 51 80 32 44
10 08 58 21 66	72 68 49 29 31	89 85 84 46 06	59 73 19 85 23	65 09 29 75 63
47 90 56 10 08	88 02 84 27 83	42 29 72 23 19	66 56 45 65 79	20 71 53 20 25
22 85 61 68 90	49 64 92 85 44	16 40 12 89 88	50 14 49 81 06	01 82 77 45 12
67 80 43 79 33	12 83 11 41 16	25 58 19 68 70	77 02 54 00 52	53 43 37 15 26
27 62 50 96 72	79 44 61 40 15	14 53 40 65 39	27 31 58 50 28	11 39 03 34 25
33 78 80 87 15	38 30 06 38 21	14 47 47 07 26	54 96 87 53 32	40 36 40 96 76
13 13 92 66 99	47 24 49 57 74	32 25 43 62 17	10 97 11 69 84	99 63 22 32 98

परिशिष्ट 'ख'

प्रश्नावली

भाग 'क' : व्यक्तिगत सूचनाएँ

अनुसूची (प्रश्न सूची) संख्या

क्षेत्र कोड

परिवार संख्या

इलाके/मोहल्ले का प्रकार

क. बहुत धनी

ख. उच्च मध्य वर्ग

ग. मध्य वर्ग

घ. बहुत गरीब (झोपड़ पट्टी)

ड.. ग्रामीण

च. शहरी

उत्तरदाता का नाम

पता

उत्तरदाता की :

क. जाति

ख. धर्म

परिवार का प्रकार

क. व्यक्तिगत

ख. संयुक्त

परिवार की रूपरेखा

क्रम संख्या	परिवार के सदस्यों के नाम	उत्तरदाता के साथ संबंध	आयु	लिंग	शिक्षा
1	उत्तरदाता	स्वयं उत्तरदाता			
2					
3					
4					
5					

रोजगार (work) स्थिति

क्रम संख्या	परिवार के सदस्यों के नाम	रोजगार सहित (E)	बेरोजगार (NE)	कार्य की प्रकृति	आय
1	उत्तरदाता				
2					
3					
4					
5					

भाग 'ख' : खाद्य सामग्री

प्रश्न: आप किराना सामग्री नियमित रूप से कहाँ से खरीदते हैं?

क. सुपर बाजार से ☐

ख. निकट की किराना दुकान से ☐

ग. अलग-अलग दुकानों से ☐

घ. अन्य (स्रोत बताएँ) ☐

प्रश्न :आप दूध नियमित रूप से कहाँ से लेते हैं?

क. मदर डेयरी बूथ ☐

ख. डी.एम.एस. बूथ ☐

ग. किराना की दुकान से पैकेटों में ☐

घ. दूध वाले से ☐

ङ. अन्य (स्रोत बताएँ) ☐

प्रश्न :आप माँस नियमित रूप से कहाँ से खरीदते हैं?

क. निकट की माँस की दुकान से ☐

ख. अन्य (स्रोत बताएँ) ☐

प्रश्न :आप मछली प्रायः कहाँ से खरीदते हैं?

क. निकट की मछली की दुकान से ☐

ख. फेरीवाले से ☐

ग. अन्य (स्रोत बताएँ) ☐

प्रश्न :आप सब्जियाँ प्रायः कहाँ से खरीदते हैं?

क. निकट की दुकान से ☐

ख. फेरीवाले से ☐

ग. अन्य (स्रोत बताएँ) ☐

भाग 'ग': कपड़ा तथा अन्य उपभोक्ता टिकाऊ वस्तुएँ

प्रश्न : क्या आप प्रायः बने-बनाए कपड़े खरीदते हैं?

क. हाँ ☐ ख. नहीं ☐ ग. कभी-कभी ☐

प्रश्न : क्या आप दर्जी से कपड़े सिलाते हैं?

क. हाँ ☐ ख. नहीं ☐

प्रश्न : आप टिकाऊ वस्तुएँ कहाँ से खरीदते हैं? जैसे — फर्नीचर, रेडियो, टी.वी., रेफ्रिजरेटर आदि। ☐

भाग 'घ': उपभोक्ता जागरूकता

प्रश्न : क्या आप खरीदते समय वस्तुओं की जाँच करते हैं?

क. हाँ ☐ ख. नहीं ☐

प्रश्न : क्या आपको कभी निम्नलिखित में मिलावट की शिकायत मिली है?

क. खाद्य सामग्री क. हाँ ☐ ख. नहीं ☐

ख. दूध क. हाँ ☐ ख. नहीं ☐

प्रश्न : यदि हाँ, तो आपने कभी निम्नलिखित के पास शिकायत दर्ज की है?

क. दुकानदार ☐

ख. मुख्य आपूर्तिकर्त्ता ☐

ग. अन्य (बताएँ) ☐

प्रश्न : क्या आपकी शिकायत विक्रेता द्वारा संतोषप्रद रूप से सुनी गई?

क. हाँ ☐ ख. नहीं ☐

प्रश्न : जब आप अन्य स्रोतों से वस्तुएँ खरीदते हैं तो क्या उनकी जाँच करते हैं?

क. हाँ ☐ ख. नहीं ☐

प्रश्न : क्या आप उपभोक्ताओं की शिकायतों की सुनावाई के लिए बनाई गई उपभोक्ता अदालतों के बारे में जानते हैं?

क. हाँ ☐ ख. नहीं ☐

प्रश्न : क्या आपने कभी उपभोक्ता अदालत में शिकायत दर्ज की है।

क. हाँ ☐ ख. नहीं ☐

परिशिष्ट 'ग'

द्वितीयक आँकड़ों के कुछ महत्त्वपूर्ण स्रोत निम्न हैं :

1. **भारत का सांख्यिकीय सारांश 2000,** जो केंद्रीय सांख्यिकीय संगठन, सांख्यिकी तथा कार्यान्वयन मंत्रालय, भारत सरकार के द्वारा प्रकाशित किया जाता है।
2. **वार्षिक योजना, 1996-97,** भारत सरकार, योजना आयोग, नई दिल्ली।
3. **भारतीय रिजर्व बैंक बुलेटिन 2000,** भारतीय रिजर्व बैंक।
4. **राष्ट्रीय लेखा सांख्यिकी 2000,** केंद्रीय सांख्यिकीय संस्थान, सांख्यिकी तथा कार्यान्वयन मंत्रालय, भारत सरकार।
5. **भारत की जनगणना 2001,** अनंतिम जनसंख्या योग, राजस्थान, जनगणना निदेशक, राजस्थान।
6. **भारतीय अर्थव्यवस्था की मासिक समीक्षा,** जनवरी 2001, सेंन्टर फॉर मानिटरिंग इंडियन इकॉनोमी प्रा.लि.।
7. **आर्थिक सर्वेक्षण 2001-2002,** भारत सरकार, वित्त मंत्रालय, आर्थिक प्रभाग।

सांख्यिकीय शब्दावली (Glossary)

समष्टि (Population) — सांख्यिकी में **समष्टि** का तात्पर्य किसी क्षेत्र की सभी व्यक्तिगत इकाइयों से संबद्ध चर या चरों के मूल्यों के कुल समुच्चय से है। यदि मूल्य किसी एक चर के संबंध में हों तो इसे **एकविचर समष्टि** कहते हैं, यदि मूल्य एक से अधिक चरों के संबंध में हों तो इसे **द्विविचर** या **बहुविचर समष्टि** कहते हैं।

प्राचल (Parameter) — **प्राचल** समष्टि का एक या एक से अधिक अनजान अचर है। उदाहरण के लिए, समष्टि का समांतर माध्य या विचरण मापांक अचर हैं।

आकलक (Estimator) — प्रतिदर्श आँकड़ों से प्राचल मूल्यों के आकलन को निकालने की विधि को **आकलक** कहते हैं। उदाहरण के लिए, प्रतिदर्श समांतर माध्य समष्टि के समांतर माध्य का एक आकलक है तथा प्रतिदर्श विचरण मापांक समष्टि के विचरण मापांक का एक आकलक है।

आकलन (Estimate) — **आकलन** किसी दिए गए प्रतिदर्श से प्राप्त आकलक का संख्यात्मक मूल्य है।

प्रतिचयन त्रुटि (Sampling Error) **प्रतिचयन त्रुटि** प्राचल के आकलन तथा इसके सही मूल्य के बीच का संख्यात्मक अंतर है।

गणना विधि (Census Method) — **गणना-विधि** आँकड़े एकत्र करने की एक विधि है, जिसमें किसी क्षेत्र की सभी व्यक्तिगत इकाइयों के प्रेक्षण लिए जाते हैं।

प्रतिदर्श विधि (Sample Method) — **प्रतिदर्श विधि** आँकड़े एकत्र करने की एक विधि है, जिसमें समष्टि के कुछ चुने हुए व्यक्तियों के प्रतिनिधि समुच्चय के प्रेक्षण लिए जाते हैं। यादृच्छिक प्रतिचयन में समष्टि की सभी व्यक्तिगत इकाइयों के चुने जाने की समान संभावना होती है। यदि समष्टि सजातीय नहीं है, परंतु इसे सजातीय स्तरों में बाँटना संभव है, तो चुनाव के लिए यादृच्छिक प्रतिचयन का प्रयोग किया जाता है।

अप्रतिचयन त्रुटियाँ (Non–sampling Error) — आँकड़ों के एकत्रीकरण में **अप्रतिचयन त्रुटियाँ** निम्नलिखित कारणों से हो सकती हैं :

क. चरों के मापन संबंधी त्रुटियों के कारण,

ख. रेकार्डिंग संबंधी त्रुटियों के कारण,

ग. अनुत्तर संबंधी त्रुटियों के कारण, आदि।

प्रश्नावली (Questionnaire) — किसी विषय-विशेष पर पूछताछ करने के लिए अन्वेषणकर्ता द्वारा तैयार की गई प्रश्नों की सूची को **प्रश्नावली** कहते हैं। उत्तरदाता को इन प्रश्नों के उत्तर देने होते हैं।

औसत (Average) — किसी चर के मूल्यों के वितरण की केंद्रीय प्रवृत्ति की माप को **औसत** कहते हैं। उदाहरण के लिए, समांतर माध्य, मध्यिका तथा बहुलक औसत हैं जो भिन्न-भिन्न प्रकार से केंद्रीय प्रवृत्ति की माप करते हैं। समांतर माध्य इस अर्थ में केंद्रीय मूल्य है क्योंकि समांतर

माध्य से (चर के मूल्यों के) संख्यात्मक विचलनों का योग शून्य होता है, मध्यिका इस अर्थ में केंद्रीय मूल्य है क्योंकि इससे अधिक मूल्यों की संख्या इससे कम मूल्यों की संख्या के बराबर होती है (यहाँ वास्तविक मूल्यों का महत्त्व नहीं होता); तथा बहुलक वह मूल्य है जिसकी आवृत्ति सबसे अधिक बार होती है।

परिक्षेपण (Dispersion) — यह केंद्रीय मूल्य से अन्य मूल्यों के परिक्षेपण या प्रकीर्ण का माप है। उदाहरण के लिए, मानक विचलन तथा माध्य विचलन समांतर माध्य से परिक्षेपण के माप हैं। परास तथा चतुर्थक विचलन परिक्षेपण के अन्य माप हैं, परंतु ये किसी मूल्य से विचलनों पर विचार नहीं करते। ये परिक्षेपण के सामान्य मापक हैं।

विभाजक मूल्य (Partition Values) — मध्यिका, चतुर्थक, दशमक, शततमक **विभाजक मूल्य** माने जाते हैं। मध्यिका चर का वह मूल्य है जो मूल्यों के समुच्चय को दो समान भागों में विभाजित करता है, चतुर्थक वे मूल्य हैं (Q_1, Q_2, Q_3) जो समुच्चय को चार बराबर भागों में, दशमक दस बराबर भागों में तथा शततमक सौ बराबर भागों में बाँटते हैं।

सहसंबंध गुणांक (Correlation Coefficient) — **सहसंबंध गुणांक** दो चरों के बीच साहचर्य का मापक है। यदि चरों के बीच रैखिक संबंध नहीं है या चरों की संख्या दो से अधिक है तो इनके बीच साहचर्य की व्याख्या करना कठिन है। यदि दो चरों के बीच रैखिक संबंध है, तो सहसंबंध का वर्ग उनके बीच साहचर्य की शक्ति की माप करता है। उदाहरण के लिए, मान लें कि सहसंबंध गुणांक 0.6 है तो $(0.6)^2 = 0.36$। इसका अर्थ है कि एक चर में 36 प्रतिशत विचरण की व्याख्या दूसरे चर के साथ इसके रैखिक संबंध द्वारा की जाती है। दूसरे शब्दों में, 64 प्रतिशत विचरण की व्याख्या इस संबंध के द्वारा नहीं होती। स्पष्टतः, $r = 0.6$ को उच्च सहसंबंध नहीं कहा जा सकता है। इसी प्रकार, मान लें कि x तथा y के बीच सहसंबंध 0.3 है तथा u तथा v के बीच 0.6 सहसंबंध है। क्या हम कह सकते हैं कि u तथा v के बीच साहचर्य की शक्ति x तथा y के बीच के साहचर्य की शक्ति से दोगुनी है? उतर होगा 'नही' क्योंकि $r^2_{x,y} = 0.09$ तथा $r^2_{u,v} = 0.36$। अतः वस्तुतः u तथा v के बीच साहचर्य की शक्ति x तथा y के बीच के साहचर्य की शक्ति से चौगुनी अधिक है। सहसंबंध गुणांक कार्य-कारण संबंध की दिशा नहीं बताता। सहसंबंध गुणांक यह नहीं बताता कि x के कारण y घटित होता है या y के कारण x।

सूचकांक (Index Number) — **सूचकांक** आर्थिक क्रिया का मापक है। यह समय के साथ कीमतों, मात्राओं तथा मूल्यों में होने वाले सापेक्ष परिवर्तनों को मापता है। उदाहरण के लिए, वर्तमान अवधि की कीमतों की तुलना आधार अवधि की कीमतों के साथ की जाती है, आदि। सूचकांक का प्रयोग नीतिगत उद्देश्यों के लिए किया जाता है।

उत्तरमाला

अध्याय 1

प्रश्न 12. 120

अध्याय 3

प्रश्न 9. i. परास = 45%

ii क. अपवर्जित विधि द्वारा

45-55	55-65	65-75	75-85	85-95	95-100
5	20	7	9	8	1

ii ख. अपवर्जित विधि द्वारा

50-55	55-60	60-65	65-70	70-75	75-80
5	7	13	5	2	8

80-85	85-90	90-95	95-100
1	5	3	1

प्रश्न 10. क. परास = 9539 रु.

अध्याय 4

प्रश्न 6. समांतर माध्य = 29.1925, मध्यिका = 28.50, बहुलक = 28

प्रश्न 7. समांतर माध्य ≈ 6.96, मध्यिका = 6.75, बहुलक ≈ 6.33

प्रश्न 8. समांतर माध्य ≈ 79.47, मध्यिका = 84, बहुलक = 84

प्रश्न 9. मध्यिका ≈ 46.74

प्रश्न 10. निम्न-चतुर्थक Q_1= 32.09 तथा उच्च-चतुर्थक Q_3= 47.70

प्रश्न 11. मध्यिका = 41.17

प्रश्न 12. $Q_1 \approx 202.38$, मध्यिका ≈ 261.90, $Q_3 \approx 390$, बहुलक ≈ 252.17

[**संतुलित असममित वितरणों** से समांतर माध्य निकालने के लिए हम निम्नलिखित **सन्निकट** संबंध का प्रयोग कर सकते हैं :

समांतर माध्य – बहुलक ≈ 3 (समांतर माध्य – मध्यिका)]

समांतर माध्य ≈ 266.765

अध्याय 7

प्रश्न 6. i. समांतर माध्य ≈ 173.16, मानक विचलन ≈ 17.27

ii. समांतर माध्य ≈ 173.72, मानक विचलन ≈ 17.52

iii. समांतर माध्य ≈ 173.67, मानक विचलन ≈ 17.96

प्रश्न 7.

	गणित	अर्थशास्त्र
परास	55%	24%
समांतर माध्य	55.07%	48.47%
मानक विचलन	15.59%	7.45%
माध्य विचलन (समांतर माध्य से)	12.59%	6.30%

परिक्षेपण के सापेक्ष मापक :

	गणित	अर्थशास्त्र
$\frac{R}{X_{max}+X_{min}}$	0.52	0.24
मानक विचलन/समांतर माध्य	0.28	0.15
माध्य विचलन/समांतर माध्य	0.23	0.13

प्रश्न 8. क. लागतों का विचरण मापांक = 144 $(Rs.)^2$, लागतों का मानक विचलन = 12 रु.
समांतर माध्य से लागतों का मानक विचलन = 9.84 रु.

ख. i. मानक विचलन / समांतर माध्य = 0.17 तथा

ii. माध्य विचलन / समांतर माध्य = 0.14 क्योंकि समांतर माध्य = 71

प्रश्न 9. क. i. $\bar{x} = 12$

ii. मानक विचलन = 3.66

iii. $\bar{x}$ से माध्य विचलन = 3.20

ख. i. $\sum (x_i - 10)^2 = 174$

ii. मध्यिका = 11, $\sum |x_i -$ मध्यिका $| =$ **30**

ग. $\sum (x_i - 10)^2 = 174$

$\sum (x_i - \bar{x})^2 = 134$

$\sum |x_i - \bar{x}| = 32$

$\sum |x_i -$ मध्यिका $| x = 30$

प्रश्न 10. विचरण मापांक = 0.01 $(\text{litre})^2$

प्रश्न 11. 2000 में विचरण का गुणांक 6 है तथा 2001 में 9 है। अतः 2001 में सापेक्ष परिक्षेपण 2000 से अधिक है। अतः 2001 में परीक्षाफलों में कोई सुधार नहीं हुआ है।

प्रश्न 12. i. समांतर माध्य = 5.5

ii. मानक विचलन ≈ 2.87

iii. समांतर माध्य से माध्य विचलन = 2.5

iv. मध्यिका से माध्य विचलन = समांतर माध्य से माध्य विचलन, क्योंकि समांतर माध्य = मध्यिका = 5.5

मानक विचलन का प्रयोग करते हुए, विंचरण का गुणांक ≈ 52%;

तथा माध्य विचलन का प्रयोग करते हुए, विचरण का गुणांक = 45.45%.

अध्याय 8

प्रश्न 11. चूँकि **Y** का विचरण मापांक शून्य (0) है, अतः हम गुणन आघूर्ण सहसंबंध के लिए इस सूत्र का प्रयोग नहीं कर सकते। हालाँकि हम देखते हैं कि प्रकीर्ण आरेख में सभी बिंदु **X**-अक्ष के समांतर सरल रेखा पर स्थित हैं। सरल रेखा की ढाल शून्य (0) है। अतः, $r = 0$.

प्रश्न 12. क. $r \approx 0.49$

ख. $r \approx 0.63$

प्रश्न 13. $r \approx 0.95$

प्रश्न 14. $r \approx 0.99$

प्रश्न 17. क. निम्नलिखित प्राप्त करने के लिए दोनों चरों को हज़ार करोड़ के निकटतम तक दर्शाएँ

निजी उपभोग व्यय (000 करोड़ रु.)	**सकल देशीय उत्पाद (000 करोड़ रु.)**
12	15
12	16
13	17
15	20
17	23
18	24
20	27
25	32
25	33
27	37

इससे प्रकीर्ण आरेख बनाने में सुविधा होगी।

ख. $r = 0.996$

प्रश्न 18. $r \approx 0.98$

अध्याय 9

प्रश्न 7. ख. 229.42%

ग. 220.55%

प्रश्न 8. 157.3, 165.2, 160.7, 175.7, 164.9 तथा 177.2

There were so many candidates who, though, did not have any confusion between average of numbers and sum of numbers, but still obtained the solution with the error No. 7 e.g.

Sum of 5 numbers = 15 x 5 = 75.

Sum of last 3 numbers = 17 x 3 = 51.

Average of first two nnumbers = 75 - 51 = 24.

Thus the above described two errors are they only common errors.

Question No. 5

Unit- Time and distance. (Relative speed)

Question- Mohan and Sohan started from Ajmer and Vijainagar respectively at 10 A.M. to see each other. Mohan and Sohan travel at speeds of 5 km. and 7 km per hour respectively. Find at what time will they meet each other and how much distance each would have travelled by then, if the distance between the two places is 60 km.

Objective- Understanding.

Processes involved in the solution :

1. Determine relative speed of Mohan and Sohan as = 12 km / hr.

2. Finds time = $\frac{60}{12}$ hrs. = 5 hrs.

3. Finds distance travelled by Mohan = 5x5 = 25 kms.

4. Finds distance travelled by Sohan = 7x5 = 35kms.

5. Determine the time of meeting =(10+5)hrs.=3 p.m.

<u>Possible errors</u> :

1. Error in writing unit of speed (one may write the unit of speed here as km only instead of km / hr.)

2. Omission of units.

3. Calculate relative speed as 7-5 = 2 km/hr.

4. One does a mistake in fiding the time e.g. One may write the time = (10 + 5)hr = Any time other than 3.00 P.M.

5. One may make an error in recalling the formula relating to distance, time and speed - In place of D=V x t. One may use D x t = V or DxV=t etc.

6. One may calculate different timings of meeting for Mohan and Sohan e.g. time for Mohan = $\frac{60}{5}$ = 12 hours and for Sohan $\frac{60}{7}$ hrs. . etc

STATISTICAL DATA (IN FIGURES)

Category	Error No.						Not Attempted.	Attempted	Irrelevant
	1	2	3	4	5	6			
I	2	2	-	3	3	1	1	50	-
II	20	3	4	4	-	6	1[illegible]	66	6
III	15	9	8	4	2	9	24	78	15
Failure	3	1	3	4	-	9	28	43	23
Total	4[illegible]	15	15	15	5	25	63	237	40
Percentage	16.9	6.3	6.3	6.3	2.1	10.5	21	79	16.9

Data in percentage :

Category	Error No.						Attempted
	1	2	3	4	5	6	
I	4	4	-	6	6	2	50
II	30.3	4.5	6.1	6.1	-	9.1	66
III	19.2	11.5	10.3	5.1	2.6	11.5	78
Failure	7	2.3	7	9.3	-	21	43

INTERPRETATION OF THE DATA :

These data show that errors No. 1 and 6 are common errors. Very few candidates omitted the unit of speed. Certainly there were many who wrote the unit of speed km. in place of km/hr. There were some who could not calculate the time of their meetingand if calculated, it was other than

3 p.m. some wrote it as 5 P.m. and a few x wrote it as 2. p.m. while there were a few who simply wrote the time as 10 + 5 = 15 hrs. etc. Usually candidates recalled the formula correctly. A few candidates obtained the relative speed as 7-5 = 2 km/hr. There were many candidates who obtained different times of their travel before they met e.g. time for Mohan = $\frac{60}{5}$ = 12 hrs. and for Sohan

$\frac{60}{7}$ hrs. etc.

Question No. 7

Unit- Equations (Simultaneous)

Objective- Application.

Question- Ten years, ago, the age of the father was five times the age of his son and twenty years hence the age of the father will be wice the age of his son. Find the age of the father.

Processes involved in the solution :

1. Suppose the present age of the son as x yrs. and consequently finds the age of the father 10 years before or vice- versa.

2. Calculates the age of the father and son, 20 years hence.

3. Forming equations.

4. Solving equations and getting present age of the father 60 years and that of his son equal to 20 years.

Possible Errors :

1. In place of multiplying the age of the son, one may multiply the age of the father in forming equation.
2. One may make a mistake in transposing terms in solving the equation.
3. One calculates the age of the son and puts it as the answer.
4. One does not consider the age before 10 years and consequently does not form the corresponding equation.

STATISTICAL DATA

Category	Error No. 1	2	3	4	Not Attempted.	Attempted.	Irrelevant.
I	-	-	-	3	35	16	2
II	2	-	-	3	56	20	10
III	1	1	-	10	59	43	16
Failure	-	-	-	7	44	27	19
Total	3	1	-	23	194	106	47
Percentage	2.8	9	-	21.7	64.7	35.3	44.3

Data in percentage :

Category	Error No.				Attempted.
	1	2	3	4	
I	-	-	-	10.75	16
II	10	-	-	10	20
III	2.3	2.3	-	25.26	43
Failure	-	-	-	25.9	27

INTERPRETATION OF THE DATA

Error No. 4 is the only common error. None of the other errors is a common error. In this problem more than 44% candidates did irrelevant work (not connected with the problem). Most of the candidates could not form the necessary equations.

Question No. 7(1) (Alternative)

Unit- Set Theory

Objective- Application

Question- If A = { 1,2,3,4 } ; B = { 2,8,6,8 } and C = { 3,4,5,6, }. then prove that A U (B U C) = (A U B) U C.

PROCESSES INVOLVED IN THE SOLUTION:

1. Finds B U C by taking all the elements of A and B and writes B U C = { 2,3,4,5,6,8 }.

2. Consequently finds $A \cup (B \cup C) = \{1,2,3,4,5,6,8\}$

3. Finds $A \cup B = \{ 1,2,3,4,6,8 \}$

4. Then finds $(A \cup B) \cup C = \{1,2,3,4,5,6,8\}$

5. Then verifies that all the elements are common in the two sets $A \cup (B \cup C)$ and $(A \cup B) \cup C$, and hence concludes that $(A \cup B) \cup C = A \cup (B \cup C)$.

Possible Errors :

1. One may not have any concept of the symbol of union and thus calculate $B \cup C$ and $A \cup B$ incorrectly.

2. One draws only the Venn. diagram for $A \cup (B \cup C)$ and $(A \cup B) \cup C$ and ends the solution there itself without further arguments.

3. One omits or writes wrong symbol for the symbol of the set i.e. one may use ( ) or [ ] in place of { } to represent a set.

4. One may not put comma between consecutive elements.

5. One may write $A \cup (B \cup C) = (A \cup B) \cup C = \{1,2,3,4,5,6,7,8\}$ i.e. in a hurry includes the element 7.

STATISTICAL DATA

Category	Error No.					Not attempted.	Attempted.	Irrevelent
	1	2	3	4	5			
I	9	-	5	1	-	22	29	4
II	11	2	10	-	-	37	39	9
III	9	1	8	1	-	63	39	8
Failure	14	-	1	-	-	44	27	8
Total	43	3	24	2	-	166	134	29
Percentage	32.1	2.2	17.9	1.5	-	55.3	44.7	21.6

INTERPRETATION OF THE DATA

Errors Nos. 1 and 3 are common errors. Most of the candidates did not have the concept of the symbol of union, hence could not find A U B and B U C and consequently A U (B U C) and (A U B) U C. Many of these candidates who attempted this part (about 18%) did not know the symbol of representing the set either they did not put any symbol or put a wrong symbol like ( ) or [ ] in place of { }

Question No. 7(ii)(Alternative)

Unit- Sets (Venn diagram).

Objective- Understanding.

Question- Represent $A \cap B$ by Venn diagram.

Processes involved in the solution :

1. Draws Venn diagram of $A \cap B$ when A and B are disjoint.

2. Draws Venn diagram of $A \cap B$ when one of these sets is a subset of the other.

3. Draws Venn diagram of $A \cap B$ when $A \cap B \neq \phi$ and none is contained in the other.

Possible errors :

1. One does not draw all the three diagrams.
2. One draws Venn diagram for $A \cup B$ in place of $A \cap B$.
3. Draws a diagram but does not shade it it show the correct portion of $A \cap B$.

STATISTICAL DATA

Category	Error No. 1	Error No. 2	Error No. 3	Not attempted.	Attempted.	Irrelevant.
I	26	-	1	24	27	1
II	23	1	-	42	34	7
III	25	-	-	64	38	5
Failure	16	-	-	48	23	6
Total	90	1	1	178	122	19

Data in terms of percentage :

Category	Error No. 1	2	3	Attempted.
I	96.3	-	3.7	27
II	67.6	2.9	-	34
III	65.0	-	-	30
Failure	69.6	-	-	23

INTERPRETATION OF THE DATA

None of the candidates drew all the three required diagrams. Almost none confused union with intersection and almost all have shaded the intersection part.

Question No. 8

Unit- Circle (Area and circumference)

Objective- Understanding.

Question- The circumference of the top of a circular table is 198 cm. Find its diameter.

Processes involved in the solution :

1. Recalls the formula $C = \pi \times d$.
2. Substitutes the values of $C = 198 = \pi \times d$.
3. Evaluates $d = \frac{198}{\pi}$ cms = 63 cm.s after substituting $= \frac{22}{7}$

Possible Errors :

1. Uses $C = \pi r^2$ instead of $C = 2\pi r$ or $C = \pi d$.
2. Calculates r and does not convert it into diametre.
3. One does not write unit.
4. One uses diameter $= \frac{C}{2\pi}$ or $d = \frac{2\pi}{C}$ or $d = \pi C$.
5. Computational Error.

STATISTICAL DATA :

Category	Error No. 1	2	3	4	5	Not attempted.	Attempted.	Irrelevant
I	1	2	-	1	-	2	49	-
II	2	1	2	9	3	7	69	6
III	4	-	2	13	2	21	81	14
Failure	2	-	1	19	3	19	52	25
Total	9	3	5	41	8	49	251	45
Percentage	3.6	1.2	2	16.3	3.2	16.3	83.7	17.9

Data in terms of percentage :

Category	Error No. 1	2	3	4	5	Not attempted.
I	2	4.1	-	2	-	49
II	2.9	1.4	2.9	13	4.3	69
III	4.9	-	2.5	16	2.5	81
Failure	3.8	-	1.9	36.5	5.7	52

INTERPRETATION OF THE DATA

The common error in this problem is the inability of recall of the correct formula. Many of these used $d = \frac{C}{2\pi}$, a very few used $d = \frac{2\pi}{C}$ or $\frac{\pi}{C}$ etc. There were only 3.6% candidates who used $C = \pi r^2$ (an incorrect formula). There were about two percent candidates who did not use unit of the diameter & circumference. A very few candidates did computational error.

Question No. 9

Unit- Cylinder.

Objective- Knowledge.

Question- The height of a cylinder is 45 cm.s and its radius is 14 cm. Find the curved and total suSface of the cylinder.

Processes Involved in the solution :

1. Using the correct formula for the curved surface i.e. $2\pi rh$.

2. One calculates the curved surface by substituting the values of the height and radius of the cylinder in the given formula and arrives at the correct result which is equal to 3960 Sq. c.m.

3. One uses the correct formula for the area of the plane surface (circular surface) which is equal to πr^2

4. Substitutes the value of the radius and gets $2\pi(14)^2$

5. Calculates the total surface i.e. 5192 sq.cm.

6. Recalls the formula for the whole surface and substitutes the correct terms and simplifies.

Possible errors :

1. One applies a wrong formula $\pi r^2 h$ instead of $2\pi rh$ for the curved surface.

2. Similarly one applies wrong formula for the area of the circular plane faces e.g. $2\pi r$ in place of πr^2 or applies any other wrong formula.

3. One may take area of one surface only instead of both with the curved surface to obtain the whole surface.

4. One may write the answer at the initial stage only where he arrives at the area of the curved surface i.e. 3960 sq.cm.

5. Computation- error.

6. One may take curved surface as the total surface.

7. One may omit to write the units of the area.

8. One may put wrong units.

STATISTICAL DATA :

Category	Error No.								Not attempted.	Attmpted.	Irrelevant
	1	2	3	4	5	6	7	8			
I	-	2	-	-	6	-	4	1	25	26	-
II	9	6	2	1	9	3	7	5	45	33	1
III	5	9	3	1	15	3	11	2	67	35	2
Failure	9	2	-	-	3	1	5	3	47	24	7
Total	23	19	5	2	33	7	27	11	182	118	10
Percentage	19.5	16.1	4.2	1.6	28	5.9	22.9	9.3	60.7	39.3	6.5

Data in terms of percentage :

Category	Error No.								Attempted.
	1	2	3	4	5	6	7	8	
I	0	7.7	0	0	23.1	0	15.4	3.9	26
II	27.3	18.2	6.1	3	27.3	9.1	21.2	15.2	33
III	14.3	25.7	8.6	2.9	42.9	8.6	31.4	5.7	35
Failure	37.5	8.3	-	-	12.5	4.2	20.8	12.5	24

INTERPRETATION OF THE DATA

1. These data show that Errors No. 1,2,5 and 7 are the common errors.

2. About 6 percent candidates were ignorant of the units and they put wrong units for the area e.g. they wrote Cm in stead of sq.cm. for the unit of area.

3. About 1/5th of the whole lot (who attempted) got confused surface area with the volume of the cylinder and about 1/6th of the whole lot got confused area of the plane surface with its circumference.

4. About 1/4th of the whole lot (who attempted) did not write the units of the area.

Most of the error committors were from second divisioners, third divisioners and failures.

Question No. 9(Alternative)

Unit- Area of rectangular paths.

Objective- Knowledge.

Question- A rectangular garden is 180 M long and 120 M wide There are two paths each 5 m wide in the middle of the garden and parallel to its length and breadth. Find the total area of the paths.

Processes involved in the solution :

1. Draws correct diagram of the two paths.

2. Finds the area of each path by multiplying the length with its width 180 x 5 sq.m., 120x5 sq.m.

3. Finds the sum of these two areas = (900 +600) sq.m.

4. Finds the area of the common path (i.e. square) = 5 x 5 = 25 sq.m.

5. Subtracts the common area from the sum of the areas of the two paths and gate = 1500-25=1475 sq.m.

Possible Errors :

1. One does not draw the path in the central portion of the field but draws it on the boundary of the field.

2. One does not take account of the area of the common path in finding the total area of the paths.

3. One does computational error.

4. One draws a wrong diagram other than described in error no. 1.

5. Omission of Units.

STATISTICAL DATA :

Category	Error No.					Not Attempted.	Attempted.	Irrelevant.
	1	2	3	4	5			
I	-	-	2	0	-	27	24	-
II	5	1	-	4	6	35	41	6
III	4	3	7	-	3	43	59	23
Failure	3	1	2	2	8	32	39	15
Total	12	5	11	6	17	137	163	44
Percentage	7.4	3.1	6.7	3.7	10.4	45.7	54.3	27

Data in terms of percentage :

Category	Error No. 1	2	3	4	5	Attempted.
I	-	-	8.3	-	-	24
II	12.2	2.4	-	9.8	14.6	41
III	6.8	5.1	11.9	-	5.1	59
Failure	7.7	2.6	5.1	5.1	20.5	39

INTERPRETATION OF THE DATA

The following are the observations from the data :

1. Omission of units is the only common error in this problem.

2. The misunderstanding of the paths at the boundary in place of centre remained a common error for second divisioners but did not remain a common error for others.

3. First and third divisioners committed error in general in the computational work.

4. About 27% of the candidates including second, third divisioners and failures did totally an irrelevant work such as found the area of the field in place of paths, or did not complete the problem, took only I step etc.

Question No. 10

Unit- Trigonometrical ratios and use of standard identities.

Objective- Understanding.

Question Prove that $\cot^2 \theta (\sec^2 \theta -1) = 1$

Processes involved in the solution :

1. Writes the value of $\sec^2 \theta - 1$ by using the formula $\sec^2 \theta = 1 + \tan^2 \theta$

2. Cancels $\cot^2 \theta$ with $\tan^2 \theta$ in theproduct to give the result 1.

Or

Alternatively.

1. Converts $\cot \theta$ and $\sec \theta$ into $\sin \theta$ and $\cos \theta$.
2. Uses $1 - \cos^2\theta = \sin^2\theta$.
3. Simplifies and gets 1.

Possible Errors :

1. One does not apply the formula $\sec^2 \theta = 1 + \tan^2\theta$ correctly or $1 - \cos^2 \theta = \sin^2 \theta$.
2. One does not recall that $\cot\theta = \frac{1}{\tan\theta}$ or $\cot\theta = \frac{\cos\theta}{\sin\theta}$
3. Computation-error.
4. Mistakes in opening the brackets.

STATISTICAL DATA

Category	Error No. 1	2	3	4	Not attempted.	Attempted.	Irrelevant.
I	7	3	5	7	22	29	1
II	9	4	4	5	49	27	11
III	12	11	6	4	82	20	6
Failure	5	4	2	-	58	13	6
Total	33	22	17	16	211	89	24
Percentage	37.1	24.7	19.1	18	70.3	29.7	26.9

Data in terms of percentage :

Category	Error No. 1	2	3	4	Attempted.
I	24.1	10.3	17.2	24.1	29
II	33.3	14.8	14.8	18.5	27
III	60	55	30	20	20
Failure	38.5	30.3	15.4	-	13

INTERPRETATION OF THE DATA

From the data it is clear that all these errors are the common errors.

It shows that a large number of candidates did not preparethis chapter of trigonometry.

Question No. 10(Alternative part)

Unit- Solution of right angled triangle(Trigonometry)

Objective- Understanding

Question- In the $\triangle ABC$, $C = 90^\circ$, $a = 6$, $b = 6\sqrt{3}$, then find the remaining elements of the triangle.

Processes involved in the solution :

1. Using Pythagoras theorem, calculate the value of the third side $C = 12$.

2. Applies the tratio $\tan A = \frac{\overline{BC}}{AC}$ for finding the value of the angle A and gets $A = 30^\circ$.

3. Calculates the value of the third angle by using $A+B+C = 180^\circ$ i.e. $A+B=90^\circ$ and gets $B = 60^\circ$.

Possible Errors :

1. One does not remember the relation between the three sides of a right angled triangle i.e. pythagoras theoram.

2. One does not remember the correct value of the t-ratio - which is used to find the value of one angle.

3. Computational error.

4. One determines one angle only or one side only.

5. One is unable to recognise the remaining elements.

6. One does not know the sides a,b,c,.

STATISTICAL DATA

Category	Error No. 1	2	3	4	5	6	Not atten-pted.	Attem-pted.	Irre-levement.
I	-	1	1	1	1	4	39	12	2
II	3	2	3	2	3	2	60	16	8
III	20	8	6	7	12	13	65	37	14
Failure	13	3	1	1	4	5	52	19	8
Total	36	14	11	11	20	24	216	84	32
Percentage	42.9	16.7	13.1	13.1	23.8	20.6	72	28	38.1

DATA IN TERMS OF PERCENTAGE

Category	Error No. 1	2	3	4	5	6	Attem-pted.
I	-	8.3	8.3	8.3	8.3	33.3	12
II	18.8	12.5	18.8	12.5	18.8	12.5	16
III	54.1	21.6	16.2	10.9	32.4	35.1	37
Failure	68.4	15.8	5.3	5.3	21.1	26.3	19

INTERPRETATION OF THE DATA

1. These data show that all the above mentioned possible errors are common errors.

2. First divisioners commit these errors very rarely.

3. About 38 percent of the candidates have done quite an irrelevent work e.g. one did not draw a

right angled triangle with right angle at C, then used some t-ratio very incorrectly (i.e. without knowing its meaning) and could not find any thing.

4. About 20 percent candidates were quite unfamiliar with the meaning of sides a,b,c e.g. one took a as AC, b as BC and c AS CA etc.

Question No. 11

Unit- Areas of rectangular fields.

Objective- Application

Question- The perimeter of a square field is 200 m. Find the length of a rectangular field whose breadth is 25 m and equal in area to that of the square.

Processes involved in the solution :

1. Writes the formula
Perimeter of a square = 4 times of the side of the square and finds the value of the side = $\frac{200}{4}$ =50m.

2. Finds the area of the square by squaring the value of its side = 50^2 = 2500 sq.m.

3. Writes the formula for the area of a rectangle: area of a rectangle = length x breadth, then takes breadth = 25 m. and Area = 2500 sq.m. and gets 2500 = 25 x 1 . (1 = length).

4. Calculates the value of the length i.e.
1 = 100 m.

Possible Errors :

1. One does not distinguish perimeter from the area.
2. One uses the incorrect formula, length = area x breadth due to wrong conception of the formula.
3. One does not discriminate square from rectangle.
4. One does not distinguish perimeter from the side of the x square.

STATISTICAL DATA :

Category	Error No. 1	2	3	4	Attempted.	Irrelevant.
I	2	-	1	1	48	6
II	16	-	-	10	67	7
III	30	-	2	6	73	23
Failure	23	1	1	18	55	24
Total	71	1	4	35	243	60
Percentage	29.2	.4	1.6	14.4	81	24.7

Data in terms of percentage :

Category	Error No. 1	2	3	4	Attempted
I	4.2	-	2.1	2.1	48
II	23.9	-	-	14.9	67
III	41.1	-	2.7	8.2	73
Failure	41.8	1.8	1.8	32.7	55

INTERPRETATION OF THE DATA

1. Data make it clear that error No. 1 and 4 are only the common errors.

2. Only one candidate used an incorrect formula like - length = Area x breadth.

3. Similarly there were only four candidates who did not discriminate a square from a rectangle.

4. Quite a large number of candidates did not know pythagoras theorem and the number of those candidates was also quite significant who did not distinguish the sides a, b and c in the triangle ABC.

Question No. 12

Unit- Pythagoras Theorem.

Objective- Knowledge.

Question- In a right angled triangle, the hypotenuse is 10 cm and one side is 6 cm. Find the length of the remaining side.

Processes involved in the solution :

1. Writes the formula

$(\text{Hypotenuse})^2 = (\text{One side})^2 + (\text{other side})^2$

2. Substitutes the values of the hypotenuse and one side and gets $10^2 = 6^2 + (\text{ other side })^2$

3. Calculates the value of the remaining side $= 8$ cm.

Possible errors :

1. One does not use the sign of equality.
2. One takes square of the side like $AC^2 = \sqrt{64} =$ 8 cm.
3. Writes area $= 10 \times 6$.
4. Applies wrong formula.
5. Computational Error.

STATISTICAL DATA

Category	Error No. 1	2	3	4	5	Attempted.	Irrelevant.
I	-	-	-	1	-	49	-
II	-	6	-	9	2	66	2
III	2	12	1	29	4	86	13
Failure	1	-	6	22	1	52	11
Total	3	18	7	61	7	253	26
Percentage	1.2	7.1	2.8	24.1	2.8	84.3	10.3

Data in terms of percentage :

Category	Error No. 1	2	3	4	5	Attempted
I	-	-	-	2	-	49
II	-	9.1	-	13.6	3	66
III	2.3	14	1.2	33.7	4.6	86
Failure	2	-	11.5	42.3	2	52

INTERPRETATION OF THE DATA

1. First divisioners did not commit any of these errors except one candidate who committed the fourth error. Thus none of these errors is a common error for first divisioners.

2. Second divisioners also committed these errors rarely. It looks that error Nos. 2 and 4 are the common errors for second divisioners.

3. The fourth error is a common error.

4. Many of the candidates could not draw even the rough sketch of the given right angled triangle. There were some candidates who even did not know the meaning of the sides, hypotenuse and one side etc. and took 10 as a side other than hypotenuse. Some drew the figure correctly but then could not apply pythagoras theorem rightly and wrote

$$10^2 + 6^2 = (\text{ other side })^2$$

etc.

It was a common fact that many wrote solution as
$(AB)^2+(AC)^2 = (BC)^2$, from the figure with A as right angle & then took the steps,

$$(AB)^2 + 6^2 = 10^2$$

or $(AB)^2 = 10^2 - 6^2 = 100-36 = 64 = 64 = 8$ Ans.

Question No. 13

Unit- Volume of a cuboid.

Objective : Application.

Question: A cistern is 4 m long, 2 m 50 cm wide and 1 m 50 cm deep. If one litre of water occupies 1000 cm^3 of space, find how many litres of water can be contained in the cistern.

Processes involved in the solution :

1. Finds the volume of the cistern = 4 x 2.5 x 1.5 = 15 m^3 15 m^3

2. Converts the volume of the cistern into cubic centimetres by multiplying it by 100^3 = 1000000 and gets the volume = 15000000 cm^3.

3. Finds the volume of water in litres by dividing with 1000 and gets the required result = 15000 litres.

Possibel Errors :

1. One may write area of the cistern = lxbxh.
2. One may not have any knowledge regarding conversions from m^3 into cm^3.
3. One can not correlate the volume of the cistern with volume of one litre of water in finding the volume of contained water in litres.
4. One shows quite ignorance about litres etc.
5. One may apply the formula volume of cistern = 2(lb+ bh+ hl).

6. Computational error.

7. One may apply volume or Area = 2(l + b) h for working of the problem.

STATISTICAL DATA

Category	Error No. 1	2	3	4	5	6	7	Attempted	Irrelevent.
I	1	15	12	15	3	6	1	48	1
II	8	35	12	14	9	8	11	69	4
III	12	25	23	21	13	9	14	86	15
Failure	11	21	19	13	11	2	7	48	13
Total	32	96	66	54	36	25	33	251	33
Percentage	12.7	38.2	26.3	21.5	14.3	10	13.1	83.7	13.1

Data in terms of percentage :

Category	Error No. 1	2	3	4	5	6	7	Attempted.
I	2.1	31.3	25	12.5	6.3	12.5	2.1	48
II	11.6	50.7	17.4	20.3	13	11.6	15.9	69
III	14	29.1	26.8	24.4	15.1	10.5	16.3	86
Failure	22.9	43.8	35.6	27.1	22.9	4.2	14.6	40

INTERPRETATION OF THE DATA.

1. Almost all the above listed errors are observed as common errors. Students frequently commit all these types of errors. Those who committed mistakes comprised of all categories of candidates.

It seems, this unit must have been taught at the end of the session and no thorough preparation of this unit was made by the candidates.

Question No. 14

Unit- Graph

Objective: Skill

Question The temperature of a patient on a certain day is given by the following table :

Time :	7 a.m.	9 a.m.	11 a.m.	1 p.m.	3 p.m.	5 p.m.
Temperature	37°C	37.2°C	38°C	39.1°C	38.5°C	37.6°C.

Processes involved in the solution :

1. Chooses appropriate scale for each of time and temperature.

2. Plots the points and draws the graph by joining the plotted points by straight lines.

3. Interprets the graph for finding the temperature at 8 a.m. and 4 p.m. and obtains temperature at 8 a.m. = 37.1°C and at 4 p.m. = 38.05°C.

Possible Errors :

1. One cannot choose appropriate scale.
2. One cannot plot points correctly.
3. One has no idea about the scale andplotting.
4. One cannot interpret the result i.e. makes mistake in interpreting the result.

5. One joins the plotted points inaccurately.

STATISTICAL DATA

Category	Error No. 1	2	3	4	5	Attem-pted.	Irr-levent.
I	9	11	7	21	1	50	-
II	25	35	26	45	3	71	2
III	37	58	59	74	3	97	-
Failure	32	36	40	41	1	65	2
Total	103	140	132	181	8	283	4
Percentage	36.4	49.5	46.6	63.9	2.8	94.3	1.4

Data in terms of percentage :

Category	Error No. 1	2	3	4	5	Attempted
I	18	22	14	42	2	50
II	35.2	49.3	~~24~~ 36.6	63.4	4.2	71
III	38.1	59.8	60.8	76.3	3.1	97
Failure	49.1	55.4	61.5	63.1	1.5	65

INTERPRETATION OF THE DATA

From the data it is clear that first, second, third and fourth errors are the common errors. A few candidates did not join the plotted points by straight lines but joined them by curved lines.

All these errors are due to the fact that students are not given enough practice of plotting the points. Most of the candidates leave this topic in choice it is why these general errors, which are due to ignorance of the subject, are taking their position among the common errors.

A consolidated list of Kernel and consequential errors occuring in section B is given in Appendix B.

CHAPTER - FIFTH

18. INTERRELATIONSHIP BETWEEN DIFFERENT UNITS, CONCLUSIONS AND SUGGESTIONS

In this section we shall try to see inter-relationship between the various (Chapters) units of the syllabus.

I. SQUARE ROOT AND CUBE ROOT

Its basic concept is derived from the index laws. When one looks into the solution of general equation

$$x^2 = a$$

in one unknown x, he has to take care of the numbers of the form $a^{1/2}$, whose square is considered to be a.

Similarly , numbers of the form $a^{1/3}$ are solutions of the equation of the form

$$x^3 = a$$

and these are those numbers whose cube is a .

Square root and cube root both are very important concepts. Both these topics have wide application in various

other Areas of thesyllabus - such as :

1. Factorisation - (Algebra)
2. Compound interest problems (Arithmetic)
3. Ratio and proportion (Arithmetic)
4. Pythagoras theorem problems (Geometry)
5. Areas of squares and circles (Geometry)
6. Volumesof Cylinder, sphere, etc.(Geometry)
7. Relation among t-ratios (Trigonometry)
8. Values of standard angles (Trigonometry).

etc.

(i) STUDY OF SQUARE ROOT AND CUBE ROOT IN FACTORIZATION :

In factorization, we come accross various expressions which involve square root of a product of two quantities, e.g. in part B of this paper if we go through the question on factorization we have in Q.1 (a) terms like $25x^4$, $4y^4$, in Q.1(b) terms like $49y^2$ etc. which are required to be written in the form $(5x^2)^2$, $(2y)^2$ and $(7y)^2$ respectively in the first step of factorization. In these cases we need square root of these terms i.e. we need to find out square root of 25 as 5, square root of x^4 as x^2, square root of $49y^2$ as 7y etc. Similarly in some other questions on factorization we are expected to know the knowledge of cube roots etc. e.g. in the factorization of

$27x^3 - 64y^6$, we are required to know the cube roots of both $27x^3$ and $64y^6$ etc.

If a candidate does not have the concept of square root and cube root or has wrong concepts of these units, he will not be able to do factorization correctly at such steps, i.e. he will not succeed to convert $25x^4$ into $(5x^2)^2$ etc. and consequently will not be able to factorize such expressions. From the collected statistics of 300 answerscripts of the candidates of all categories of the Exam. 1972, we observe that about 19 per ent candidates of could not factorize Q 1(a) on account of the wrong concept of square root : These 19 percent candidates included 6 percent first divisioners, 19 percent second divisioners, 18 percent third divisioners and 35 percent failures. Had this concept of square root of a product of two terms been clear, it would have certainly enabled them to factorize it correctly and consequently would have improved their result and also the pass percentage in the subject at the Board. Similarly in Q 1(b) this lack of knowledge of square root has led about 21 percent of the candidates to failure in factorization including about 14 percent first divisioners, 5 percent second divisioners, 23 percent third divisioners and 42 percent failures.

(ii) SQUARE ROOT OR CUBE ROOT IN COMPOUND INTEREST PROBLEM:

If there is a problem in which it is said that a sum P becomes the amount A in two or three years, find the rate of compound interest, one will be required to use the formula

$$A = P(1 + \frac{i}{100})^2 \quad \text{or} \quad A = P\ (1 + \frac{i}{100})^3 \text{ etc.}$$

and then to find i it is necessary to find the square root or cube root of A/P. It shows that the knowledge of cube root or square root is needed in solving various problems of compound interest, annuities, etc.

(iii) IN RATIO AND PROPORTION :

If there are problems of the type. ' Find the mean propo tional of a and b ', then we have to find the mean proportional = $\sqrt{ab}$, i.e. here also we require the concept of square root.

(iv) IN PYTHAGORAS THEOREM PROBLEMS :

In all the numerical problems on y pythagoras theorem, one needs the knowledge of square root e.g. In Q. 12 of Part II of this paper, we are given hypotenuse = 10 cm and one side = 6 cm. and we are asked to find the value of the third side, It's solution is

$$10^2 = 6^2 + (\text{ other side })^2$$

$$(\text{other side})^2 = 10^2 - 6^2 = 100 - 36$$
$$= 64.$$
$$\therefore \text{other side} = \sqrt{64} = 8 \text{ cm.}$$

Thus here also we require the concept of square root. The collected statistics show that due to lack of knowledge of this concept about 3 percent candidates could not find the correct answer of this problem.

A similar was the situation in Q. 10 (alternative part).

(v) IN AREAS OF SQUARES AND CIRCLES :

In both the cases if area of square or circle is given and it's side or radius is required, we always need to use the concept of square root.

(vi) IN VOLUMES OF SPHERE OR CYLINDER :

In the problem on these units if the volume is given and the radius or base radius is required, we have need to use cube root or square root etc.

(vii) IN TRIGONOMETRY (IN RELATION AMONG T- RATIOS)

In the problems where value of a particular t-ratio is given and the values of other t-ratios are required, we need to use the concept of square root e.g. in Q 16

(Part A).

(viii) IN TRIGONOMETRY – ∈

(In height and distances problems) etc.

II. PERCENTAGE

Percentage is another important concept. It also has applications in various other units of the syllabus such as :

1. Simple interest,
2. Compound interest,
3. Profit and loss,
4. Partnership.

etc.

In simple x interest the rate of interest is always given in terms of percentage e.g. at the rate of 6 percent per annum. etc. In compound interest too the rate of interest is given in percentebe. In profit, and loss, the profit and loss are also given in percentabe. In partnership, it is sometimes said that of the total sum the shares of the partners are x percent, y percent, z percent, respectively etc. Thus it is basic to learn percentage before we proceed to learn these other units. Due to lack of knowledge of percentabe about 10.2% candidates committed error of using unitary law incorrectly.

III. SIMPLE INTEREST :

The concept of simple interest is basic for compound interest problems and increase in population problems e.g. In Q. No. 3 of Part A of this paper, it is required to find out the compound interest on Rs. 1,000/- for four years at the rate of 10 percent annually and in the Q. No. 4 of Part B, the present population of a city is given and it is required to find out the population after two years when it increase at the rate of 10 percent annually.

It shows that one can only do these problems if one has the concept of simple interest.

Because of lack of theknowledge of simple interest about 2 percent candidates could not attempt Q.No. 4 of Part B correctly.

IV. RATIO AND PROPORTION :

The concept of ratio and proportion is basic in the other units like division into proportional parts, partnership and trigonometry.

Q. Nos. 7 and 8 of Part A are based on ratio and proportion. Question Nos. 16 , 19 of Part A and Q. No. 10 of Part B are from trigonometry and partnerships etc. All these

questions require basic knowledge of ratio and proportion, without which none can be solved.

V. FACTORS :

Factors are basic in simple, simple simultaneous and quadratic equations which are indirectly or directly used in various problems of geometry and trigonometry (including mensuration). Since there was no problem on these units in this question paper, it could not be illustrated with the help of daxx data.

VI. LOGARITHMS :

It is used to find out values of various types of numerical expressions involving multiplication, division involution and evolution operations and is thus required whenever such expressions are involved in problems, may be from Algebra, Geometry or Trigonometry. Hence its knowledge is basic in all these fields. Since there was no problem on the application of this unit in this question paper, it could not be illustrated with data.

VII. AREA OF A RECTANGLE :

The concept of area of a rectangle is basic for the following units :

i) area of a parallelogram;

ii) area of a triangle;

iii) area of ~~kxx~~ a trapezium;

iv) area of irregular fields ;

v) pythagoras theorem;

vi) area of four walls;

vii) surface area of cuboids;

viii) surface area of prisms ;

ix) problems of carpeting and flooring;

x) area of a circle; and

xi) Area of curved surface of a cylinder.

We know that :

i) the rectangles and parallelograms drawn on the same base or on equal bases and between the same parallels are always equal and thus derive that area of parallelogram = base length x perpendicular distance between two parallels including the base;

ii) a triangle and a rectangle or a parallelogram are situated on the same base and between the same parallels then area of that triangle is equal to half of thearea of that rectnangle;

iii) area of trapezium is equal to sum of the areas of two triangle obtained by drawing a diagonal;

iv) the area of the a square drawn on the hypotenuse of a right angled triangle is equal to thesum of

the areas of the squares drawn on its other two sides (Pythagoras theorem);

v) all the walls of a room are rectangles and thus area of four walls is the sum of the areas of the four rectangles;

vi) the surface area of cuboids and prisms also depend on the areas of rectangles and triangles involved; similar is the situation in the problems of carpeting or flooring of a room ;

vii) area of a circle is derived from the area of triangles in limiting case ;

viii) surface area of a cylinder is exactly an area of a rectangle.

All the above mentioned facts show that the concept of the area of a rectangle is basic for all the above mentioned units which are in one or the other manner used in the problems of daily life.

In this question paper I the analysis of Q.No. 17 and 18 of part A and Q. Nos. 9, 10, 11, 12 illustrate p this point clearly.

VIII. TRIGONOMETRICAL RATIOS :

Every one who has read trigonometry knows that

trigonometrical ratios are basic in the complete theory of trigonometry and in the problems on the solution of triangles or we can say in the problems on height and distances.

For want of the knowledge of this topic, about 10.1% candidates made error in solving the right angled triangle of Q. No. 10 (alternative part).

2. CONCLUSIONS AND SUGGESTION :

The present trend of setting question Papers and evaluating answer books is worth praising. But it some times fail to achieve its objective. The setting of Part A of the question paper is not an easy job. It requires a lot of practice to set multiple choice questions. Stems (distractors) formed are sometimes useless, and non-functional. If a paper has too many questions of this nature, it certainly improves the result of the examination and from no corner any hue and cry is heard but the paper does not make a proper evaluation.

I, therefore, suggest that (i) in Part A, the number of multiple choice questions should be reduced from 16 to 10 or 12 and short- answer questions may be increased from 6 to 8 or 10; (ii) the setters may be asked to give the explanation for settin a question under a

particular objective ; (iii) the setter may also be asked to give the expectationof errors in forming the distractors with incorrect results.

All these measures will enable the Board to get good questions set in the question papers. For short answer questions of Part A, some space may be provided to write one or two steps to arrive at the answer.

In chapter V we have established interrelationship between different units. If the teachers take care of the fundamental mistakes committed in square roots, cube roots, recentages and sums which lead to consequential errors in other areas of mathematics as exemplified in the chapter II and III and the Appendix A and B, it would definitely lead to improvement the class-room teaching and the examination results of the Board.

I suggest that all the common errors which are obtained in Chapter Third and Fourth alongwith this discussion on interrelationship of errors between different units of the syllabus should be brought into the handsof all the mathematics' teachers and mathematics' book-writers with a view to giving them advantage of the same while teaching and writing text books.

APPENDIX - A

A LIST OF KERNEL AND CONSEQUENTIAL ERRORS :

OBSERVED IN DIFFERENT UNITS OF SECTION

The symbol KE denotes Kernel error and CE denotes consequential error CE 2.3 means consequential error No. 3 of Kernel error No. 2 of that unit.

UNIT - SQUARE ROOT

Q.No. 1 Part - A

KE.1 Error of not converting an integer + a fraction in the form of $\frac{m}{n}$, where m and n are integers.

KE 2- Error of ignorance of the knowledge of finding square root of the numbers of the form $\frac{m}{n}$ as such uses $\sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{b}$ as in Q 1 part A of this paper.

CE 1.1 In question No. 1 part A one finds square root of $1\frac{9}{16}$ as square root of 1 x square root of $\frac{9}{16}$ and gets $1 \times \frac{3}{4}$ for $\sqrt{1\frac{9}{16}}$

CE 1.2 One finds $\sqrt{1\frac{9}{16}}$ = Square root of ($1 + \frac{9}{16}$)

= Square root of 1 + square root of $\frac{9}{16}$

and gets it = $1 + \frac{3}{4} = 1\frac{3}{4}$ or $\frac{7}{4}$ etc.

Data showing the occurrance of these errors in this investigations are as :

CE 1.1	6%
CE 1.2	$19\frac{1}{3}$ %
KE 2	$6\frac{2}{3}$ %

UNIT PERCENTAGE

Q.No.2 Part A

KE 1. Error of ignorance of the knowledge of the meaning of percentage - i.e. one does not know $47\% = \frac{47}{100}$ in Q. No. 2 of part A.

	IN CONSEQUENTIAL ERRORS	Percentage of occurrance.
CE 1.1	Taking $47\% = \frac{47}{1000} = 0.047$	$12\frac{1}{3}$
CE 1.2	Taking $47\% = \frac{47}{10} = 4.7$	10
CE 1.3	Taking $47\% = 47$	4

UNIT COMPOUND INTEREST

Q. No. 3 Part A

KE 1 Ignorance of the knowledge of the meaning of the terms AMOUNT, PRINCIPAL, INTEREST etc.

KE 2. Ignorance of the knowledge of the method of determining compound interest.

CE 1.1 One determines amount, when interest is asked as some candidates have done in Q. No. 3 part A

Weightage of this error in the data = 43%.

CE 1.2 For any given Principal one finds interest = Amount - Rs. 1/- . As in Q. No. 3 part A, candidates have answered, compound interest on Rs. 1000/- for four years @ 10% = Rs. $\left\{1000\left(1+\frac{10}{100}\right)^4-1\right\}$

CE 2.1 Finds interest = Principal - $\left(1+\frac{R}{100}\right)^n$

where R is the rate of interest and n is the number of (intervals) years.

Weightage of this error in the data = $3\frac{2}{3}$ %.

UNIT PROFIT AND LOSS

Q.No. 4 Part A

KE 1 Error of ignorance of the meaning of the words cost price, selling price and profit or loss etc. in Hindi equivalent words क्रय मूल्य, विक्रय मूल्य और लाभ हानि आदि)

KE 2 Error of ignorance of the fact that profit or loss is calculated on cost price only.

KE 3 Error in using Unitary law incorrectly.

KE 4. Error due to Hastiness.

CE 1.1 When cost price is to be determined, one determines total profit or when total profit is asked, one determines selling price only. Some times selling price is given, profit percentage is given and cost price is wanted one uses these data as if he calculates profit on selling price or calculates cost price as

$$\text{Selling price} \times \frac{(100 + R)}{100}$$

where R is the rate of profit percentage.

This error can also be considered in consequence of KE 3. This error had weightage in the data = $33\frac{1}{3}$%

CE 2.1 Uses profit = S.P. x $\frac{\text{rate percent of profit}}{100}$

or

Uses cost price = $\frac{\text{S.P.x rate percent of profit}}{100}$

It actually is in consequence of KE .1 & KE 2 both. The weightage of this error in the data = $7\frac{1}{3}$%

CE 3.1 = C.E. 1.1

CE 4.1 Due to hastiness one may calculate the cost price by taking the data given for profit percentage as loss percentage and consequently arrives at

$$\text{C.P.} = \text{S.P.} \times \frac{100}{(100 - R)}$$

The weightage of the in error in the data

$= 2\frac{2}{3}\%$

UNIT RELATIVE SPEED

Q.No. 5 Part A

KE 1 One confuses the notion of relative speed when they proceed in the same direction with the one when they proceed in opposite direction.

KE 2 Relative speed involves the operations of + or -, but due to ignorance of the subject matter i.e. of the knowledge of relative speed, he uses any other operation such as or x etc.

CE 1.1 When one has to determine the relative speed of two objects moving in opposite direction, he actually determines the relative speed ax by the formula which is used when they proceed in the same direction or vice versa .

It was committed by $20\frac{1}{3}\%$ of the candidates in this sample.

CE 2.1 One finds relative speed as speed of one x speed of second.

It has weightage in the data = 2%

CE 3.1 One finds relative speed as

Speed of one speed of other.

It occupied a weightage of $\frac{2}{3}$ % in the data.

UN II WORK AND TIME

Q.No. 6 Part A

When it is given that A can do a work in x days and B in y days and their one day's joint work is asked None of the expected errors can be a Kernel error.

?

?

KE 1. In such problems, one can only does a mistake of the form that for finding one day's joint work one first adds the number of days in which A does the work with the number of days in which B does the work andthen finds the joint one day's work as $\frac{1}{x+y}$ and it is none of the expected errors of this problem.

KE 2. One can proceed to find one day's joint work = one days work of A x one days work of B. It is also none of the expected errors of this problem.

KE 3. One may not understand the meaning of one day's joint work and consequently find's either oneday's work of A or one day's work of B.

CE 3.1 Finds one day's joint work = one day's work of A
The weightage of this error in the data = 1%

CE 3.2 Finds one day's joint work = one days work of B
The weightage of error in the data = 1/3%

UNIT WORK, TIME AND WAGES

Q. No. 7 In this problem, it is given that two persons Ram and Moti can do a piece of work in 20 days and 25 days respectively. If they do the work jointly and receive a sum of Rs. 90/-, what shall be the Ram's share.

ME 1. One can not correlate the data and consequently divides the amount equally to them or finds the shares in any other manner.

ME 2. One finds theproper shares but can not relate them to their owners, consequently finds Ram's share as the share of Moti.

CE 1.1 Shows Ram's share = Rs. 45/-
Weightage of this error in the data 13%

CE 1.2 Shows Ram's share = Rs. 90/= - 20/= Rs. 70/-
i.e. the amount Rs. 90/- is diminished by the no. of days in which Ram completes the work.
Weightage of this error in the data = 2.1/3%

CE 2.1 Finds share of R_{gn} = Rs. 40/- which is actually the share of Moti.

Weightage of this error in the data = 46.2/3%

UNIT RATIO AND PROPORTION

Q. No. 8 Part A

When A : B :: 6: 7 and B : C :: 14 : 17 are given and A: B : C is required -

KE 1 One may lack the knowledge of the fact that in both the relations $\frac{A}{x} = \frac{B}{y}$ and $\frac{B}{u} = \frac{C}{v}$, one has to make the denominator's of B equal, consequently he may ignore one of the data of B and writes A , B: C from rest of the data. or he may add the data of B in determining A : B : C.

CE 1.1 Writes A : B : C::6 : 7: 17

Weightage of this error in the data = 15.1/3%

CE 1.2 Writes A : B : C:: 6 : 14: 17

Weightage of this error in the data = $17\frac{2}{3}$ %

CE 1.3 Writes : B: C: = 6:(7+14): 17 = 6:21:17

Weightage of this error in the data = $17\frac{2}{3}$ %

UNIT AVERAGE

Q. No.9 Part A

KE 1 One may not have any idea of the fact that average

of certain numbers always lies between the least and greatest of them.

CE 1.1 One may determine it as common difference of the given numbers if they are given in Arithmetic progression.

Weightage to this error in the data = 2/3%

CE 1.2 One may add all thenumbers and divide by 2.

Weightage to this error in the data = $3\frac{2}{3}$%.

CE 1.3 One may only add them to find the average .

Weightaged to this error in the data = 4.1/3%

UNIT **SQUARE [illegible] ROOT AND CUBE ROOT**

Q.No. 10 Part A

In this question square root of cube root of 64 is asked.

KE 1 One may confuse in this problem in understanding that whether it is a question on square root or a question on cube root.

KE 2. One may square root of cube root 64 as the number divided by its cube root.

CE 1.1 Finds cube root only and answer as 4.

Weightage to this error in the data = 28.1/3%

CE 1.2 Finds square root only and answer as 8

Weightage to this error in the data = 20%

CE 2.1 Finds the answer : $\frac{64}{\text{cube root of } 64} = \frac{64}{4} = 16$

Weightage to this error in the data = $1\frac{1}{3}$ %.

UNIT LOGARITHMS

Q. No.11 Part A

In this question recall of the formula

$\log m^n = n \log m$

is asked.

KE 1 Error of ignorance of the meaning of m^n and the knowledge of the formula $\log m^n = n \log m$.

KE 2. Error due to hastiness.

CE 1.1 Writes $\log m^n = \log \frac{m}{n}$

Weightage to this error in the data = 20%

CE 1.2 Writes it = log mn

Weightage to this error in the data = $1\frac{1}{3}$ %

CE 1.3 Writes it = m log n

Weightage to this error in the data = $14\frac{1}{3}$ %

UNIT SETS

Q.No.12 Part A

In this question symbol of void (empty) set is to be recognised.

KE 1 Since 0 in the number system stands of nothing when associated with the things. One may think {0} to stand for a set having no element i.e. void or empty set.

KE 2 Similarly in set theory { } & ϕ are symbols introduced for void set, one may not feel any difference between {ϕ} and ϕ or { } and {{ }} whereas and {ϕ} and {{ }} each is a single ton set.

CE 1.1 Choose the symbol {0} for void set.
The weightage to this error in the data = 5%

CE 2.1 Choose the symbol {ϕ} for void set.
The weightage oto this error in the data = $34\frac{1}{3}$ %

CE 2.2 Choose the symbol {{ }} for void set.
The weightage to this error in the data = $8\frac{1}{3}$%

UNIT SIMPLE LINEAR EQUATIONS

Q.No.13 Part A
In this question solution $\frac{5x}{6} = 10$ is asked.

KE 1 One does not verify the equation from the given answers and makes an error of ignoring ~~an~~ one of the figures in the equation which leads him to an incorrecced answer.

KE 2 Computational error.

CE 1.1 Ignores 6 and finds $x = \frac{10}{5} = 2$

Weightage to this error in the data = $7\frac{2}{3}$ %

CE 1.2 Ignores 5 and finds x = 10 x 6 = 60

Weightage to this error in the data = 5%

CE 2.1 Finds x = 10 x 5 by doing computational error.

Weightage to this error in the data = $4\frac{1}{3}$ %

UNIT TRIGONOMETRICAL RATIOS

Q.No.14 Part A

It is a question on recall of the formula of inverse of sin θ = cosec θ

KE 1 Error of the ignorance of the meaning of the word inverse of sinθ

CE 1.1 Since $\sin^2\theta + \cos^2\theta = 1$ takes cosθ as inverse of sinθ

Weightage to this error in the data = $24\frac{1}{3}$ %

CE 1.2 Since sin θ and sec θ both have their first letter as s and he has read relations between sinθ & cosec θ and cos θ and sec θ, confuses and takes sec θ Sin θ = 1 and thus answers inverse of Sin θ = Sec θ

Weightage to this error in the data = 10%

CE 1.3 Writes inverse of sinθ = tan θ

Weightage to this error in the data = $5\frac{1}{3}$ %

UNIT PROBLEMS IN EQUATIONS

Q.No.15 Part A

The question reads as " Square of a number x exceeds it by 42 ". The equation involving x is -

KE 1. Ignorance of the meaning of the word exceed mathematically.

KE 1.1 Finds the equation as $x^2 + x = 42$ as if he understands that exceeds means addition

Weightage to this error in data = 9⅓%

CE 1.2 One finds the equation as $x - x^2 = 42$ as if one unerstands exceeds meaning subtraced

Weightage to this error in data = 28⅔ %

CE 1.3 One understands meaning of exceeds by division and hence obtains the required equation in the form $x^2 \div x = 42$

Weightage to this error in the data = 2⅓ %

UNIT THE CALCULATION OF THE VALUES OF t-ratios FROM THE GIVEN VALUE OF ANOTHER t-ratio

Q.No.15 Pt. A

Given Cos $\theta = 4/5$, Find tan θ

KE -1 Error of ignorance of thedefinition of tan θ

KE-2 Error of the ignorance of the relation between tan θ and cos θ

CE 1.1 One takes tan θ as sin θ and finds its value = 3/5

CE 1.2 One takes tan θ as cot θ and finds its value = 4/3

CE 2.1 One takes tan θ as cos θ it self or think

= CE 1.3 tan θ = cos θ and choose its value = 4/5

UNIT RELATION BETWEEN AREA OF A TRIANGLE AND A RECTANGLE SITUATED ON THE SAME BASE AND BETWEEN THE SAME PARALLELS

Q.No.17 Part A

In this question the relation between area of rectangle and area of triangle is asked when they are situated between the same parallels andon the same base.

KE 1 Error of ignorance this relation.

CE 1.1 Takes Area of rectangle = Area of triangle
Weightage to this error in data = $16\frac{2}{3}$ %

CE 1.2 Takes Area of rectangle = Thrice the area of triangle
Weightage to this x error = 16%

CE 1.3 Takes area of rectangle = 4xarea of the triangle
Weightage to this error = ~~10.2/3%~~ $10\frac{2}{3}$ %

UNIT AREA OF WALLS OF A ROOM

Q.No.18 Part X A

In this question area of smaller wall of the room 12m x 10 m x 6 m, is asked.

KE 1 Error of not recognising the data.

KE -2 Error due to lack of concentration while reading the problem.

CE 1.1 Considers 10m and 6 m as length and breadth and 12m as height and consequently writes answer as 12 x 6 = 72 sq. m.

Weightaged to this error in data = 17.2/3%

CE 1.2 Confuses smaller wall with the floor and chooses the required area = 12 x 10 = 120 sq.m.

Weightage to this error in data = 11%

CE 2.1 One does not read the problem properly and considers it as a question on four walls and chooses answer as 2 (12 + 10) x 6 sq.m.

Weightage to this error in data = $60\frac{2}{3}$ %

APPENDIX - D

A LIST OF KERNEL ERRORS AND CONSEQUENTIAL ERRORS

AS DISCERNED FROM SECTION 'E'

N.B. W (C E 1.2) stands for " the weightage of the error CE 1.2 in the data ".

UNIT FACTORS

QUESTION NO. 1(a) FACTORIZE $25x^4 + 20x^2y^2 + 4y^4$

KE 1 Error of language due to lack of information and practice of it.

KE 2 Error of the ignorance of simple monomial factors i.e. of the ignorance of the knowledge of obtaining square roots of product of two terms.

KE 3 Error due to wrong concept of the formula $(a + b)^2 = a^2 + b^2 + 2ab$

CE 1.1 One does not use the sign of equality between two different steps in the solution.

W (CE 1.1) = 52.7%

CE 1.2 One misses to write + between two terms of an expression

W (CE 1.2) = 13.6%

~~xxxxxx xxxxxxxxxxxxxxxxxxxxxxxxxx~~

CE 2.1 Writes $4y^2 = 2(y^2)^2 = (2y)^2$ etc.

U (CE 2.1) = 19%.

CE 2.2 One is unable to write the factors.

U (CE 2.2) = 26%

CE 3.1 One writes $25x^4 + 20x^2y^2 + 4y^4 = (5x^2)^2 + (2y^2)^2$

U (CE 3.1) = 4%

CE 3.2 One writes the answer as $(5x + 2y)^2$ or simply $5x^2 + 2y^2$ etc.

U (CE 3.2) = 6.5%

Q.No.1(b) FACTORIZE = $x^2 - 49y^2$

KE 1 and KE 2 are the same as given for Q.No.1(a)

CE 1.1 Same as given for Q No. 1 (a)

CE 1.2 Same as given for Q . No. 1(a)

U (CE 1.2) = 1.3%

CE 1.3 One writes " taking square root, we have "-

U (CE 1.3) = .4%

CE 2.1 Same as given for question No. 1 (a)

U (CE 2.1) = 21%

CE 2.2 Sameas given for question no. 1(b)

U (CE 2.2) = 18%

QUESTION NO. 1(c) FACTORIZE x^3+x+x^2+1

KE 1 Same as given for Q. No. 1 (a)

KE 2 Error of ignorance of the knowledge of factorization by grouping.

CE 1.1 Same as given for Q.No. 1 (a)

U (CE 1.1) = 65.5%

CE 1.2 Same as given for Q.No. 1 (a)

U (CE 1.2) = 23.7%

CE 2.1 One can not group the terms

U (CE 2.1) = 11.3%

CE 2.2 After grouping one can factorize each group of terms

U (CE 2.2) = 17%

CE 2.3 One cannot write factors.

U (CE 2.3) = 15%

QUESTION NO. 1 (c) FACTORIZE $2 + 5x + 3x^2$

KE 1 Same as given for Q.No. 1 (a)

KE 2 Error of ignorance of the knowledge of factorization of trinomials.

CE 1.1 Same as given for Q.No. 1 (a)

U (CE 1.1) = 65.6%

CE 1.2 Same as given for Q.No. 1 (a)

U (CE 1.2) = 25.6%

CE 2.1 One cannot split up 5x into 2x + 3x

U (CE 2.1) = 20%

CE 2.2 — One after spliting up 5x into 2 x + 3x , cannot group the terms correctly

W (CE 2.2) = 23.4%

CE 2.3 — One cannot factorize each group of terms

W (CE 2.3) 30.5%

CE 2.4 — One cannot write the factors

W (CE 2.4) = 20%

UNIT — CUBE ROOT — QUESTION NO. 2

Find the cube root of $2\frac{10}{27}$

KE 1. — Error of ignorance of the knowledge of symbol and method of finding cube root.

KE 2 — Error of language.

CE 1.1 — One does not use symbol of cube root

W (CE 1.1) = 83.6%

CE 1.2 — One does not discrionimate cube root from suare root

W (CE 1.2) = 8.4%

CE 1.3 — One finds $3\sqrt[3]{\frac{8}{5}} = 3\sqrt[3]{\frac{1}{5}}$

CE 2.1 — One writes the solution as

$\frac{64}{27} = \frac{4}{3} = 1\frac{1}{3}$ etc.

W (CE 2.1) = 20.0%

UNIT SIMPLE INTEREST QUESTION NO. 3

At what rate percent of interest, the sum of Rs. 600/- will amount to Rs. 720/- in 3 years 6 months.

KE 1 Error of the ignorance of the concept of interest or rate of interest.

KE 2 Error of the ignorance of the symbols of the formula or the formula itself.

KE 3 Error of the ignorance of the fact that in calculating interest, the time is to be taken in one unit only i.e. either in years or in months.

KE 4 Error in simplification.

CE 1.1 One does not know that interest is always calculated on Principal

U (CE 1.1) = 2.8%

CE 1.2 One takes the interest on Rs. 100/- for 7/2years as the rate percent of interest

U (CE 1.2) = 1.9%

CE 2.1 One knows formula but cannot substitute the data

U (CE 2.1) = 6.5%

CE 2.2 One could not use the formula, tried for unitary law but could not succeed to arrive at the correct result :

W (CE 2.2) = 10.2%

CE 3.1 Makes error in converting 3 years 6 months into years

W (CE 3.1) = .9%

CE 4.1 Error in simplification

W (CE 4.1) = 19%

QUESTION NO. 3 (ALTERNATIVE PART)

Find the simple interest on Rs. 780/- for 2 years 1 months at the rate of 75 p. per hundred per month.

KE 1 Error of the ignorance of the formula or the symbols used in the formula.

KE 2 Error in simplification and in converting Rs 75p/100 per month into Rs. 3/4 /100 per year or in converting 2 years 1 months into 25 months.

CE 1.1 One is unable to substitute the data in the formula

W (CE 1.1) = 4.9%

CE 1.2 One is unable to recall the formula

U (CE 1.2)=17.8%

CE 2.1 Error in simplification

(CE 2.1) = 22.7%

CE 2.2 Error in converting 75p/100 per month into per hundred per year

U (CE 2.2) = 28.8%

UNIT COMPOUND INTEREST QUESTION NO. 4

The population of a city is 106400. If the population increase by 10% every year. What will be the population after two years.

KE 1 Error of the ignorance of the concept of compound interest.

KE 2 Error of confusing between increased population and increase in population.

KE 3 Error in using the data in finding increase in population after one or two years and error in simplification.

CE 1.1 Finds increase in population in two years as a problem on simple interest

U (CU 1.1) = 29.9%

CE 2.1 FINDS INCREASE in @population in two years and ¤ writes it as the answer for thepopulation ¤ after two years

U (CU 2.1) = 5%

CE 3.1 Calculates increase in population in correctly as

x $\frac{106400 \times 160}{110}$ etc

U (CE 3.1) = 2.1%

CE 3.2 Simplification error.

U (CE 3.2) = 19.2%

UNIT AVERAGE QUESTION NO.5

Average of five numbers=15,

Average of last three = 17,

find the average of first two.

KE 1 Error of ignorance of the concept of average or wrong concept of average.

KE 2 Error of confusing sum of numbers with average of numbers.

KE 3 Computational error.

CE 1,1 Takes average of first two = Average of five - average of lastthree

U (CE 1.1) = 1.9%

CE 1.2 Takes average = $\frac{15 + 17}{2}$ = 16

U (CE 1.2) = 3%

CE 1.3 Takes average = $\frac{17 - 15}{2}$ = 1

U (CE 1.3) = 1.1%

CE 2.1 Writes average of first two = 75-51=24 etc.

U(CE 2.1) = 27.3%

CE 2.2 Average of numbers = sum of numbers 1 15x5 etc.

U (CE 2.2) = 16.0%

CE 3.1 Computational error

U (CE 3.1) = 5%

CE 3.2 Finds sum of first two 1 sum of five + sum of last three

U (CE 3.2) = 4.1%

UNIT TIME AND DISTANCE Q.NO. 6

(Relative speed)

Mohan and Sohan started from Ajmer and Vijainagar respectively at 10 A.M. to see each other. Mohan and Sohan travel at speeds of 5 kms. and 7 km per hour respectively. Find at what time will they meet each other and how much distance each would have travelled by them, if the distance between the places is 60 kms.

KE 1 Error of ignorance of the unit of speed

KE 2 Error of confusing this relative speed with the one when they travel in the same direction.

KE 3 Error in recalling the formula D = V x t.

KE 4 Error in understanding that they meet after travelling for the same time or error in finding the time of meeting.

CE 1.1 Writes wrong units of speed

U (CE 1.1) = 16.0%

CE 1.2 Does not write unit of speed.

U (CE 1.2) = 6.3%

CE 2.1 Finds relative speed = 7-5 = 2 km/hr.

U (CE 2.1) = 6.3%

CE 3.1 Takes D x t = V or D x V = t etc.

U(CE 3.1) = 2.1%

CE 4.1 Finds time of meeting other than 3 hr or 15 hrs.

U(CE 4.1) = 6.3%

CE 4.2 Finds different times for both

U (CE 4.2) = 10.6%

UNIT SIMULTANEOUS EQUATIONS QUESTION NO. 7

(Problems)

Ten years ago, the age of the father was five times the age of his son and twenty years hence the age of the father will be twice the age of his son. Find the age of the father.

KE 1 Error of not understanding the question, thus can not from the required equations or wrongly form the equations.

KE 2 Computational mistake due to lack of concentration.

Please see the remaining portion of this section on page 133.

UNIT SETS Q.NO. 7(ii)

Represent $A \cap B$ by Venn diagram

KE 1 One may have an error of not drawing the diagrams of all possible cases

N (KE 1) = 75%

UNIT CIRCLE (AREA AND CIRCUMFERENCE)

Question No. 8

The circumference of the top of a circular table is 198 cm. Find its diameter.

KE 1 Error of confusing area with circumference.

KE 2 Error due to lack of concentration.

KE 3 Error in recalling the formula for circumference.

CE 1.1 Uses $\pi r^2 = 198$

U (CE 1.1) = 3.6%

CE 2.1 Finds r and answer its value as answer

U (CE 2.1) = 1.2%

CE 2.2 Omit unit

U (CE 2.2) = 2%

CE 2.3 Computational error

U (CE 2.3) = 3.2%

CE 3.1 Uses $d = \frac{c}{2\pi}$ or $d = \frac{2\pi}{c}$ or $d = \pi c$

U (CE 3.1) = 16.3%

UNIT CYLINDER Q.NO. 9

The height of a cylinder is 45 cm and its radius is 14 cm. Find the curved and total surface of the cylinder.

KE 1 Error in recalling formula for circular plane faces and curved surface etc.

KE 2 Error in recalling for mula for total surface.

KE 3 Error of omitting units or of writing wrong units or of computation.

~~CE 1.1 Formation of incorrect equations~~

~~U (CE 1.1) = 2.8%~~

Please see remaining portion of this Section on page 134 marked 'X'

C.E. 1.1. Formation of incorrect equations

$\omega(CE\ 1.1) = 2.8\%$

CE 1.2 Cannot from an equation

$\omega\ (CE\ 1.2) = 21.7\%$

CE 2.1 Mistakes in simplification

$\omega\ (CE\ 2.1) = 0\%$

CE 2.2 Mistakes in answering i.e. shown son's age as father's age

$\omega(CE\ 2.2) = 0\%$

UNIT SETS Q. No. 7(B) (ALTERNATIVE)

If $A = \{1, 2, 3, 4\}$, $B = \{2, 4, 6, 8\}$ and $C = \{3, 4, 5, 6\}$ then prove that $A \cup (B \cup C) = (A \cup B) \cup C$

KE 1 Error of ignorance of the concept of Union

KE 2 Error of the ignorance of the symbols used to represent sets.

KE 3 Computational error.

KE 4 Error of ignoring calculations and of using only venn diagrams to prove the result.

CE 1.1 Finds $A \cup B$, $B \cup C$ incorrectly showing ignorance of the concept of union

$\omega\ (CE\ 1.1) = 32.1\%$

CE 2.1 Error in symbol of representing set

$\omega\ (CE\ .1) = 17.9\%$

CE 2.2 Error of not putting comma between consequtive elements in the set.

U (CE 2.2) = 1.5%

CE 3.1 Writes A U(B U C) = (~~# # #~~) (A U B) U C

UC = {1,2,3,4,5,6,7,8}

U (CE 3.1) = 0%

CE 4.1 Uses venn diagrams to prove theresult

U (CE 4.1) = 2.2%

Please see remaining information of this section on page 13

* CE 1.1 Takes curved surface $\pi r^2 h$.

U (CE 1.1) = 19.8%

CE 1.2 Takes plane surface area = $2 \pi r$

U (CE 1.2) = 16.1%

CE 2.1 Takes total surface = $2 \pi rh + \pi r^2$

U (CE 2.1) = 4.4%

CE 2.2 Takes total surface = curved surface

U (CE 2.2) = 5.9%

CE 2.3 Does not compute total surface as he can not recall its formula

U (CE 2.3) = 1.8%

CE 3.1 Computation error

U (CE 3.1) = 20%

CE 3.2 Wrong units

U (CE 3.2) = 9.3%

CE 3.3 Omission of units

U (CE 3.3) = 22.5%

UNIT AREA OF RECTANGULAR PATHS

T.No. 9 (Alternative)

A rectangular garden is 180 m long and 120 m wide. There are two paths each 5 m wide in the middle of the garden and parallel to its length and breadth. Find the total area of the paths.

KE 1 Error of misunderstanding the position of paths.

KE 2 Error in understanding about common path.

KE 3 Error of units and of computation.

CE 1.1 One draws the path inside and along the boundary i.e. not in the middle of the garden

U (CE 1.1) = 7.4%

CE 1.2 One draws the two paths either outside & along the boundary of the garden or at any other incorrect position

U (CE 1.2) = 3.7%

CE 2.1 One does not take account the area of common path which is repeated in finding the total area of the paths

U (CE 2.1) = 3.1%

CE 3.1 One does error in computation

ψ (CE 3.1) = [illegible].7%

CE 3.2 One omits units or writes wrong units

ψ (CE 3.2) = 10.4%

UNIT TRIGONOMETRICAL RATIOS AND USE OF STANDARD IDENTITIES.

Q.No. 10 Prove that $\cot^2\theta\ (\sec^2\theta - 1) = 1$

KE 1 Error of ignorance of the knowledge of standard identities.

KE 2 Error in opening brackets, or computation error.

CE 1.1 One does not use the formula $\sec^2\theta = 1+\tan^2\theta$ or $1-\cos^2\theta = \sin^2\theta$ correctly

ψ (CE 1.1) = 37.1%

CE 1.2 One does not recall the formula $\cot\theta = \frac{1}{\tan\theta}$ or $\cot\theta = \frac{\cos\theta}{\sin\theta}$

ψ (CE 1.2) = 24.7%

CE 2.1 Computation error

ψ (CE 2.1) = 19.1%

CE 2.2 Mistake of opening brackets.

ψ (CE 2.2) = 18%

UN IT SOLUTION OF RIGHT ANGLED TRIANGLE

Q.No. 10 (Alternative)

In the $\triangle ABC$, $C = 60^{0}$, $a = 5$, $b = 5\sqrt{3}$, the find the remaining elements of the triangle.

KE 1 Error of ignorance of the knowledge of the sides a,b,,c and the pythagoras theorem.

KE 2 Error in recognising the remaining elements.

KE 3 Error of ignorance of the knowledge of trigonometrical ratios of standard angles.

U (KE 3) = 16.7%

KE 4 Computation error.

U (KE 4) = 83.1%

CE 1.1 One does not know Pythagoras theorem

U (CE 1.1) = 42.9%

CE 1.2 One does not recognise the sides a,b,c

U (CE 1.2) = 28.6%

CE 2.1 One determines one angle only or one side only

U (CE 2.1) = 13.1%

CE 2.2 One is unable to recognise remaining elements

U (CE 2.2) = 23.8%

UNIT PERIMETER AND AREA OF RECTANGULAR FIELDS:

Q.No.11 The perimeter of a square field is 200 m. Find the length of a rectangular field whose breadth is 25 m, and area equal to that the square.

KE 1 Error of ignorance of the concept of the perimeter, Area of rectangle, rectangle and a square.

CE 1.1 Takes perimeter of the square = its area

U (CE 1.1) = 25.2%

CE 1.2 Takes perimeter square = its one side

U (CE 1.2) = 14.4%

CE 1.3 One uses length = Area x breadth

U (CE 1.3) = .4%

CE 1.4 One takes rectangle and square the same

U (CE 1.4) = 1.6%

UNIT PYTHAGORAS THEOREM

Q.No.12 In a right angled triangle, the hypotenuse is 10 cm and one side is 6 cm. Find the length of the remaining side.

KE 1 Error of the ignorance of the knowledge of Pythagoras theorem.

KE 2 Computational error or language error.

CE 1.1 Finds Area = 10 x 6 unnecessarily and incorrectly inplace of the remaining side

U (CE 1.1) = 2.0%

CE 1.2 Applies any other wrong formul to find the remaining side

W (CE 1.2) = 24.1%

CE 2.1 Computation error

W (CE 2.1) = 2.8%

CE 2.2 Does not put sign of equality between two different steps.

W (CE 2.2) = 1.2%

CE 2.3 Writes $AD^2 = \sqrt{64} = 8$ cm etc.

W (CE 2.3) = 7.1%

UNIT VOLUME OF A CUBOID

Q.No. 13 A cistern is 4 m long, 2m 5 cm wide and 1 m 50 cm deep. If one litre of water occupies 1000 cm^3 of space. Find how many litres of water can be contained in the cistern.

KE 1 Error of ignorance of the concept of volume of a cuboid. Consequently there remains confusion in the formula for the volume of cuboid.

KE 2 Error of ignorance of the knowledge of units.

KE 3 Computation error

W (KE 3) = 10%

CE 1.1 Instead of volume, takes area of cistern = abc i.e. l x b x h

U (CE 1.1) = 12.7%

CE 1.2 Takes volume of cuboid = 2(lb + bh + hl)

U (CE 1.2) = 14.3%

CE 1.3 Takes volume of cuboid = 2(l+b)h

U (CE 1.3) = 13.1%

CE 2.1 One can not convert m^3 into cm^3 or cm^3 into m^3.

U (CE 2.1) = 30.2%

CE 2.2 One cannot correlate the volume of one litre of water with volume of cuboid.

U (CE 2.2) = 26.3%

CE 2.3 One has no idea about litre etc.

U (CE 2.3) = 21.5%

UNIT GRAPH

Q.No. 14

KE 1 Error of ignorance of the knowledge of selecting proper scale

U (KE 1) 1 36.4%

KE -2 Error of ignorance of the knowledge of plotting of points with respect to the chosen scale and drawing graph etc.

APPENDIX - C

SIGNIFICANT ERRORS OF SECTION B OF ESSAY AND SHORT ANSWERTYPE QUESTIONS.

	Error Nos.
Q.No. 1(a)	1,2,3,4
Q.No. 1(b)	1,3,4,
Q.No. 1(c)	1,2,3,4,5,
Q.No. 1(d)	1,2,3,4,5,6,
Q.No. 2	1,4
Q.No. 3	1,3,4
Q.No. 3(alt)	2,3,4,5
Q.No. 4	1,4
Q.No. 5	1 and 7
Q.No. 6	1 and 6
Q.No. 7	4
Q.No. 7(i) (alt.)	1 and 3
Q.No. 7(ii) (alt.)	1
Q.No. 8	4
Q. No. 9	1,2,5 and 7.
Q.No. 9(alt.)	5
Q.No. 10	1,2,3,4
Q.No.10(alt)	all the six
Q.No. 11	1 and 4.

Q.No. 12 3 4

Q.No. 13 all seven

Q.No. 14 1,2,3 and 4.

I suggest that these errors should be sent to all the teachers teaching the subject, so that they may use it to improve their teaching. These errors can also be used by the book writers and paper setters in improving their work.

****